



# **MODERN METROLOGY CONCERNS**

Edited by **Luigi Cocco**

---

# **MODERN METROLOGY CONCERNS**

---

Edited by **Luigi Cocco**

## **Modern Metrology Concerns**

Edited by Luigi Cocco

### **Published by InTech**

Janeza Trdine 9, 51000 Rijeka, Croatia

### **Copyright © 2012 InTech**

All chapters are Open Access distributed under the Creative Commons Attribution 3.0 license, which allows users to download, copy and build upon published articles even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications. After this work has been published by InTech, authors have the right to republish it, in whole or part, in any publication of which they are the author, and to make other personal use of the work. Any republication, referencing or personal use of the work must explicitly identify the original source.

As for readers, this license allows users to download, copy and build upon published chapters even for commercial purposes, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

### **Notice**

Statements and opinions expressed in the chapters are those of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

**Publishing Process Manager** Anja Filipovic

**Technical Editor** Teodora Smiljanic

**Cover Designer** InTech Design Team

First published May, 2012

Printed in Croatia

A free online edition of this book is available at [www.intechopen.com](http://www.intechopen.com)  
Additional hard copies can be obtained from [orders@intechopen.com](mailto:orders@intechopen.com)

Modern Metrology Concerns, Edited by Luigi Cocco

p. cm.

ISBN 978-953-51-0584-8



---

# Contents

---

**Preface IX**

**Section 1 Introduction 1**

- Chapter 1 **Performance Parameters  
Measurement and Stochastic Processes 3**  
Miroslav Svítek

**Section 2 Length, Distance and Surface 27**

- Chapter 2 **Length Metrology and Calibration Systems 29**  
Saulius Kaušinis, Albinas Kasparaitis and Aurimas Jakštas

- Chapter 3 **Dimensional Metrology and  
Frequency Sweeping Interferometry 51**  
Alexandre Cabral, José Manuel Rebordão and Manuel Abreu

- Chapter 4 **Nano-Accuracy Surface  
Figure Metrology of Precision Optics 77**  
Shinan Qian and Peter Takacs

**Section 3 Voltage, Current and Frequency 115**

- Chapter 5 **Josephson Junctions for  
Present and Next Generation Voltage Metrology 117**  
Vincenzo Lacquaniti and Andrea Soso

- Chapter 6 **Measurement of Harmonic Voltage,  
Current, and Power at Industrial Frequency 143**  
Lu Zuliang

- Chapter 7 **RF and Microwave Power Sensor  
Calibration by Direct Comparison Transfer 175**  
Yueyan Shan and Xiaohai Cui

- Chapter 8 **Fast Method for Frequency Measurement by Rational Approximations with Application in Mechatronics** 201  
Daniel Hernandez-Balbuena, Oleg Sergiyenko,  
Patricia L. A. Rosas-Méndez, Vera Tyrsa and Moises Rivas-Lopez
- Section 4 Optics** 221
- Chapter 9 **Optical Radiation Measurement** 223  
A. A. Gaertner
- Chapter 10 **Optical Measurements:  
Polarization and Coherence of Light Fields** 263  
O. V. Angelsky, P. V. Polyanskii, I. I. Mokhun,  
C. Yu. Zenkova, H. V. Bogatyryova, Ch. V. Felde,  
V. T. Bachinskiy, T. M. Boichuk and A. G. Ushenko
- Chapter 11 **Digital Demodulation of Interferometric Signals** 317  
Tristan J. Tayag and R. Collins Watson
- Chapter 12 **Attosecond Metrology and Spectroscopy** 333  
Martin Schultze and Reinhard Kienberger
- Chapter 13 **Measurement of Ultrashort Optical Pulses** 349  
Yuqiang Deng, Qing Sun, Shiyong Cao,  
Jing Yu, Ching-yue Wang and Zhigang Zhang
- Section 5 Time and Relativity** 363
- Chapter 14 **Time Measurement and Earth Rotation** 365  
Victor Tissen, Alexander Tolstikov and Galina Simonova
- Chapter 15 **Relativistic Metrology: From Earth to Astrophysics** 395  
Luca Lusanna
- Section 6 Biology and Medicine** 417
- Chapter 16 **Development of Metrology for Modern Biology** 419  
Sang-Ryoul Park, Jun-Hyuk Choi and Ji-Seon Jeong
- Chapter 17 **Dimensional Measurement for Dentistry** 447  
Reisha Rafeek, Kevin Seymour and Lifong Zou



---

# Preface

---

After four centuries a quote from Galileo Galilei is still the most effective way to introduce a Metrology book: “Measure what is measurable, and make measurable what is not so”. Every human activity needs an objective measurement and every global process should be supported by a common standard.

There are several handbooks, tutorials and journals that focus on the basic concepts of metrology, the uncertainty theory and the international reference standards. Moreover, the engineering literature is awash with measurement methods for each specific field. The aim of this book is different: it offers the most recent developments in some key sectors, its goal is to represent the state of art in the main metrological areas. This book is a combination of twenty scientific publications, written in a rigorous and eloquent style by academic authors from five continents. Not only will you find highlights of important technological contributions, you will also learn about some current critical issues and relative proposals to improve the future steps. The script is directed both to students/researchers and professionals.

The volume is organized in six sections:

- First section (Introduction) includes a dissertation about the system performance parameters. Basic statistical indicators are described via formal definitions and illustrative examples in order to link measurement practice with probability laws.
- Second section (Length, Distance and Surface) is dedicated to dimensional measurements. The first chapter includes a synopsis and analysis of literature and existing scientific and technical solutions of precision 1D length calibration. It also addresses the problems of research and development of an interferometer-controlled comparator that is operated in non-ideal environmental conditions and enables one to trace the calibration of line scale of up to  $L \leq 3.5$  m long to the wavelength standard. Second chapter is devoted to absolute distance metrology, it shows that Frequency Sweep Interferometer is a particularly versatile technique for absolute dimensional metrology, allowing different approaches for different applications with different needs, both in terms of range and accuracy. Third chapter presents some techniques about Precision optical surface figure measurement; the main emphasis is on Nano-accuracy surface measurement, which is the most important aspect for optical systems development of telescopes, microscopes, cameras and imaging or focusing systems.



- Third section (Voltage, Current and Frequency) examines some electrical/electronic measurements. First chapter describes the modern voltage metrology which is based on Josephson junctions; the extension of DC Josephson voltage standard to AC and the most advanced solutions in terms of integration, power dissipation and noise immunity. Second chapter focuses on power measurements under non-sinusoidal waveform conditions; harmonic voltage, harmonic current, and harmonic power at industrial frequency, with aspects such as its principle, setup, uncertainty evaluation, and experimental validation also covered. Third chapter provides calibration methods of RF and microwave power sensor with system setup, modeling, equations, and analyses in different representations, traceability and measurement uncertainty evaluations. Fourth chapter presents a fast method for frequency measurement by rational approximations with application in Mechatronics; theoretic operation principle, experimental and simulation results for the frequency measurement method are presented in a wide frequency range.
  
- Fourth section (Optics) provides direction for the measurement of optical radiation under different configurations and for several applications. First chapter is an accurate introduction to optical metrology, it involves all necessary bases: international terminology, predominant quantities, geometrical relationship, standards, uncertainties, photometers, spectrometers and light sources. Second chapter is devoted to consideration of metrological aspects of intrinsically interconnected characteristics of light fields, such as intensity, polarization and coherence; further it represents both experimental and data processing techniques leading to high-sensitive and reliable diagnostics of biological tissues. Third chapter is oriented toward recent digital signal processing techniques used in interferometry; it reviews digital demodulation algorithms for interferometric metrology systems and presents a new (unpublished) digital demodulation algorithm. Fourth chapter explains phenomena about light sources with short and ultra-short pulse duration; it discusses the parameters of the driving laser: pulse duration, pulse energy, photon energy, and the number of attosecond bursts (single pulses or a train). Fifth chapter reports some characterization techniques about femtosecond/picosecond optical pulses measurement (autocorrelation and Wavelet transform).
  
- Fifth section (Time and Relativity) is based on tools/methods for measuring time and relativistic metrology. First chapter gives an historical review, the algorithms to define the orbit of the spacecraft by ground-based measurements and the techniques used to forecast Earth rotation parameters. Second chapter canvases relativistic metrology in terms of standards and conventions used to describe physical phenomena on the Earth's surface and space physics near the Earth and in the Solar System.

- Sixth section (Biology and Medicine) carries the reader through advances of metrology applications in biotechnology and dentistry. First chapter is focused on technical issues in metrology for biology; it describes recent methods for quantitation of deoxyribose nucleic acids and characterization techniques for various biologic measurands like proteins, cells, microorganisms, and glycans. Second chapter examines main methods for 3D measurement of free form oral surfaces; specifically, it considers the width of the crown preparation margin, the angulation of the margin and the taper angle formed when a tooth is prepared for a crown.

Special thanks to authors for their concerted efforts and to my mother for her continuous encouragement.

**Dr. Luigi Cocco**  
Automobili Lamborghini S.p.A.  
L/Q-2 Supply Quality  
Electronic Parts Responsible  
Italy



# **Section 1**

## **Introduction**



# Performance Parameters Measurement and Stochastic Processes

Miroslav Svátek  
Czech Technical University in Prague  
Faculty of Transportation Sciences, Prague  
Czech Republic

## 1. Introduction

In this chapter, the emphasis is on the analysis of performance parameters, assigned to different system components, functions or processes, together with the methodology of their statistical assessment based on limited measured samples, fulfilling the condition of normal distribution. The practical examples of GNSS (Global Navigation Satellite Systems) applications assessment and certification are presented.

## 2. Definition of performance parameters

The methodology for the definition and measurement of following individual system parameters is being developed within the frame of complex system's assessment. The basic performance parameters can be defined as follows:

- **Accuracy** is the degree of conformance between system true parameters and its measured values that can be defined as the probability

$$P(|p_i - p_{m,i}| \leq \varepsilon_1) \geq \gamma_1 \quad (1)$$

that the difference between the required system parameter  $p_i$  and the measured parameter  $p_{m,i}$  will not exceed the value  $\varepsilon_1$  on probability level  $\gamma_1$  where this definition is applicable for all  $N$  system parameters  $p_1, p_2, \dots, p_N$ .

- **Reliability** is the ability to perform a required function (process) under given conditions for a given time interval that can be defined as the probability

$$P(|\bar{v}_t - \bar{v}_{m,t}| \leq \varepsilon_2) \geq \gamma_2, t \in \langle 0, T \rangle \quad (2)$$

that the difference between required system functions (processes) represented by parameters  $\bar{v}_t$  and the vector of measured parameters  $\bar{v}_{m,t}$  will not exceed the value  $\varepsilon_2$  on probability level  $\gamma_2$  in each time interval  $t$  from the interval  $\langle 0, T \rangle$ .

- **Availability** is the ability to perform required functions (processes) at the initialization (triggering) of the intended operation that can be defined as the probability

$$P(|q_i - q_{m,i}| \leq \varepsilon_3) \geq \gamma_3 \quad (3)$$

that the difference between the required rate<sup>1</sup> of successful performing of the function  $i$  (process  $i$ )  $q_i$  and the measured  $q_{m,i}$  will not exceed the value  $\varepsilon_3$  at the probability level  $\gamma_3$ .

- **Continuity** is the ability to perform required functions (processes) without non-scheduled interruption during the intended operation that can be defined as the probability

$$P(|r_i - r_{m,i}| \leq \varepsilon_4) \geq \gamma_4 \quad (4)$$

that the difference between the required rate of successful performing of the function  $i$  (process  $i$ ) without interruption  $r_i$  and the measured  $r_{m,i}$  will not exceed the value  $\varepsilon_4$  at the probability level  $\gamma_4$ .

- **Integrity** is the ability to provide timely and valid alerts to the user, when a system must not be used for the intended operation, that can be defined as the probability

$$P(|S_i - S_{m,i}| \leq \varepsilon_5) \geq \gamma_5 \quad (5)$$

that the difference between the required rate of successful performing of the alert limit (AL)  $i$  not later than predefined time to alert (TTA)  $S_i$  and the measured  $S_{m,i}$  will not exceed the value  $\varepsilon_5$  on the probability level  $\gamma_5$ .

- **Safety** can also be covered among the performance parameters, but the risk analysis and the risk classification must be done beforehand with a knowledge of the system environment and potential risk, and then the safety can be defined as the probability

$$P(|W_i - W_{m,i}| \leq \varepsilon_6) \geq \gamma_6 \quad (6)$$

that the difference between the required rate of  $i$  risk situations  $W_i$  and the measured ones  $W_{m,i}$  will not exceed the value  $\varepsilon_6$  on the probability level  $\gamma_6$ .

A substantial part of the system parameters analysis is represented by a decomposition of system parameters into individual sub-systems of the telematic chain. One part of the analysis is the establishment of requirements on individual functions and information linkage so that the whole telematic chain can comply with the above defined system parameters.

The completed decomposition of system parameters will enable the development of a methodology for a follow-up analysis of telematic chains according to various criteria (optimisation of the information transfer between a mobile unit and a processing centre, maximum use of the existing information and telecommunication infrastructure, etc.).

The following communication performance parameters quantify the quality of telecommunication service [16]:

---

<sup>1</sup>  $q_{m,i} = \frac{Q_i}{Q}$  where  $Q_i$  is the number of successful experiments (successful performing of the function

$i$ , successful performing of the process  $i$ ) and  $Q$  is the number of all experiments (both successful and unsuccessful).

- **Availability** – (i) Service Activation Time, (ii) Mean Time to Restore (MTTR), (iii) Mean Time between Failure (MTBF) and (iv) Virtual Connection Availability
- **Delay** - is an accumulative parameter effected by (i) Interfaces Rates, (ii) Frame Size, and (iii) Load / Congestion of all active nodes (switches) in the line
- **Packet/Frames Loss**
- **Security**

Performance indicators described for communications applications must be transformed into telematic performance indicators structure, and vice versa. Such transformation allows for a system synthesis.

Transformation matrix construction is dependent on detailed communication solution and its integration into telematic system. Probability of each phenomena appearance in the context of other processes is not deeply evaluated in the introductory period. Each telematic element is consequently evaluated in several steps, based on a detailed analysis of the particular telematic and communications configuration and its appearance probability in the context of the whole system performance. This approach represents a subsequent iterative process, managed with the goal of reaching the stage where all minor indicators (relations) are eliminated, and the major indicators are identified under the condition that relevant telematic performance indicators are kept within a given tolerance range.

### 3. Quality of measured performance parameters

In this chapter, unified approach applicable for all above mentioned performance parameters [18] will be introduced.

- **Absolute measuring error** ( $\mu_a$ ) is the difference between a measured value and the real value or the accepted reference

$$\mu_a = x_d - x_s \quad (7)$$

$x_d$  - measured dynamic value

$x_s$  - corresponding real value or accepted reference

- **Relative measuring error** ( $\mu_r$ ) is the absolute measuring error divided by a true value given by

$$\mu_r = \frac{x_d - x_s}{x_s} \quad (8)$$

- **Accuracy** ( $\delta$ ) of a measuring system is the range around the real value in which the actual measured value must lie. The measurement system is said to have accuracy  $\delta$  if:

$$x_s - \delta \leq x_d \leq x_s + \delta \quad (9)$$

or straightforwardly:

$$-\delta \leq \mu_a \leq +\delta \quad (10)$$

Accuracy is often expressed as a relative value in  $\pm \delta \%$ .



- **Reliability ( $1-\alpha$ ) of a measuring system** is the minimal probability of a chance that a measuring error  $\mu_a$  lies within the accuracy interval  $[-\delta, \delta]$  :

$$(1 - \alpha) \leq P(|\mu_a| \leq \delta) \quad (11)$$

where  $P(\cdot)$  means the probability value.

- **Error probability ( $\alpha$ ) of a measuring system** is the probability that a measured value lies further from the actual value than the accuracy:

$$\alpha \geq P(|\mu_a| > \delta) \quad (12)$$

The reliability of measuring system is often controlled by the end-user of the measurement system while error probability is generally assessed by the International Organization for Legal Metrology (OIML).

- **Dependability ( $\beta$ ) of an acceptance test** is the probability that - on the basis of the sample - a correct judgment is given on the accuracy and reliability of the tested system:

$$P(\alpha \leq P(-\delta < \mu_a < \delta)) \geq \beta \quad (13)$$

The desired dependability determines the size of the sample; the higher the sample, the higher the dependability of the judgment.

## 4. Estimation of performance parameters

### 4.1 Tests of normality

With regard to [12] normal distribution will be expected, because using different kinds of statistics, such as order statistics (distribution independent) for small sample sizes, typical for performance parameters, the result may be fairly imprecise. Testing normality is important in the performance parameters procedure, because in analyses containing a lot of data this data is required to be at least approximately normally distributed. Furthermore, the confidence limits assessment requires the assumption of normality. Several kinds of normality tests are available, such as [1]:

- Pearson test (Chi-Square Goodness-of-Fit Test)
- Kolmogorov-Smirnov test
- Anderson-Darling and Cramer-von Mises test

All the above mentioned tests for normality are based on the *empirical distribution function* (EDF) and are often referred to as *EDF tests*. The empirical distribution function is defined for a set of  $n$  independent observations  $X_1, X_2, \dots, X_n$  with a common distribution function  $F(x)$ . Under the null hypothesis,  $F(x)$  is the normal distribution. Denote the observations ordered from the smallest to the largest as  $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ . The empirical distribution function,  $F_n(x)$ , is defined as

$$\begin{aligned} F_0(x) &= 0, & x < X_{(1)} \\ F_i(x) &= \frac{i}{n}, & X_{(i)} \leq x < X_{(i+1)}, \quad i = 1, \dots, n-1 \\ F_n(x) &= 1, & X_{(n)} \leq x \end{aligned} \quad (14)$$

Note that  $F_n(x)$  is a step function that takes a step of height  $1/n$  at each observation. This function estimates the distribution function  $F(x)$ . At any value  $x$ ,  $F_n(x)$  is the proportion of observations less than or equal to  $x$ , while  $F(x)$  is the probability of an observation less than or equal to  $x$ . *EDF statistics* measure the discrepancy between  $F_n(x)$  and  $F(x)$ .

In the following part the *Pearson test* (Chi-Square Goodness-of-Fit Test) will be introduced as a practical example of EDF tests. The chi-square goodness-of-fit statistic  $\chi_q^2$  for a fitted parametric distribution is computed as follows:

$$\chi_q^2 = \sum_{i=1}^L \frac{(m_i - n \cdot p_i)^2}{n \cdot p_i} \quad (15)$$

where  $L$  is the number of histogram intervals,  $m_i$  is the observed percentage in  $i$ -th histogram interval,  $n$  is the number of observations,  $p_i$  is the probability of  $i$ -th histogram interval computed by means of theoretical distribution. The degree of freedom for the chi-square test  $\chi^2$  is equal to  $L-r-1$ , where  $r$  is parameters number of theoretical distribution (in case of normal distribution  $r=2$ ).

#### 4.2 Estimation of measuring system's accuracy, reliability and dependability

Let us assume we have a normally distributed set of  $n$  measurements of performance parameters  $\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n}$  (absolute error between prescribed and measured parameters as defined in (7)).

If the mean value or a standard deviation is not known we can estimate both the mean value  $\bar{\mu}_a$  and standard deviation  $s_a$  from the measured data as follows:

$$\begin{aligned} \bar{\mu}_a &= \frac{1}{n} \sum_{i=1}^n \mu_{a,i} \\ s_a &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\mu_{a,i} - \bar{\mu}_a)^2} \end{aligned} \quad (16)$$

Let  $n$  be non-negative integer,  $\alpha, \beta$  are given real numbers ( $0 < \alpha, \beta < 1$ ) and let  $\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n}, \mu_{a,y}$  be  $n+1$  independent identically distributed random variables.

*Tolerance limits*  $L = L(\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n})$  and  $U = U(\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n})$  are defined as values so that the probability is equal to  $\beta$  that the limits include at least a proportion  $(1-\alpha)$  of the population. It means that such limits  $L$  and  $U$  satisfy:

$$P\{P(L < \mu_{a,y} < U) \geq 1 - \alpha\} = \beta \quad (17)$$

A *confidence interval* covers population parameters with a stated confidence. The *tolerance interval* covers a fixed proportion of the population with a stated confidence. *Confidence limits* are limits within which we expect a given population parameter, such as the mean, to lie. *Statistical tolerance limits* are limits which we expect a stated proportion of the population to lie within.

For the purpose of this chapter we will present only results derived under the following assumptions:

- $\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n}, \mu_{a,y}$  are  $n+1$  independent normally distributed random variables with the same mean  $\mu_0$  and variance  $\sigma_0^2$  (equivalently  $\mu_{a,1}, \mu_{a,2}, \dots, \mu_{a,n}, \mu_{a,y}$  is a random sample of size  $n+1$  from the normal distribution with mean  $\mu_0$  and variance  $\sigma_0^2$ ).
- The symmetry about the mean or its estimation is required.
- The tolerance limits are restricted to the simple form  $\bar{\mu}_a - k \cdot s_a$  and  $\bar{\mu}_a + k \cdot s_a$ , where  $k$  is a so called *tolerance factor*,  $\bar{\mu}_a$  and  $s_a$  are *sample mean and sample standard deviations*, respectively, given by (16).

Under the above given assumptions, the condition (17) can be rewritten as follows:

$$P\left\{\Phi\left(\frac{U - \mu_0}{\sigma_0}\right) - \Phi\left(\frac{L - \mu_0}{\sigma_0}\right) \geq 1 - \alpha\right\} = \beta \quad (18)$$

where  $\Phi$  is the distribution function of the normal distribution with mean zero and standard deviation equal to one:

$$\Phi(u) = \frac{1}{\sqrt{2 \cdot \pi}} \int_{-\infty}^u e^{-\frac{1}{2}t^2} dt \quad (19)$$

The solution of the problem to construct tolerance limits depend on the level of knowledge of the normal distribution, i.e., on the level of knowledge of mean deviation  $\bar{\mu}_a$  and standard deviation  $s_a$ .

In the following part the *accuracy, reliability and dependability of the measuring system* will be mathematically derived for a known mean value and standard deviation, for a known mean value and unknown standard deviation, and for both an unknown mean value and standard deviation.

### Known mean value and standard deviation

We can start with the equation [3]:

$$P\left\{P\left[\mu_0 - z_{(1-\alpha/2)} \cdot \sigma_0 \leq \mu_{a,y} \leq \mu_0 + z_{(1-\alpha/2)} \cdot \sigma_0\right] \geq (1-\alpha)\right\} = 1 \quad (20)$$

where  $\mu_{a,y}$  is the measured value,  $\mu_0, \sigma_0$  are known mean value and standard deviation and  $z_{(1-\alpha/2)}$  is a percentile of normal distribution (e.g. for  $\alpha = 0.05$  we can find in statistical table  $z_{0,975} = 1.96$ ).

Based on (20) we can decide that measuring system's accuracy  $\delta = z_{(1-\alpha/2)} \cdot \sigma_0$  is guaranteed with measuring system's reliability  $(1-\alpha)$ . Because the mean value and standard deviation are known, the measuring system's dependability is equal to  $\beta = 1$ .

### Known standard deviation and unknown mean value

Now we expect that the mean value is estimated according to (16). Then we can write the equation [3]:

$$P\left\{P\left[\bar{\mu}_a - k \cdot \sigma_0 \leq \mu_{a,y} \leq \bar{\mu}_a + k \cdot \sigma_0\right] \geq (1 - \alpha)\right\} = \beta \quad (21)$$

where  $\sigma_0^2$  is the known variance and  $k$  is computed from the following equation:

$$\Phi\left(\frac{z_{(1+\beta)/2}}{\sqrt{n}} + k\right) - \Phi\left(\frac{z_{(1+\beta)/2}}{\sqrt{n}} - k\right) = 1 - \alpha \quad (22)$$

where the function  $\Phi(u)$  was defined in (19) and sample  $\bar{\mu}_a$  computed according to (16).

Based on the equation (22) we can say that for the predefined values of measuring system's reliability  $(1 - \alpha)$  and dependability  $\beta$  and the number of measurements  $n$  the accuracy of measuring system will be

$$\delta = \left( z_{1-\frac{\alpha}{2}} + \frac{1}{\sqrt{n}} \cdot z_{\frac{(1+\beta)}{2}} \right) \cdot \sigma_0 \quad (23)$$

### Known mean value and unknown standard deviation

For a known mean value and unknown standard deviation we can write the equation:

$$P\left\{P\left[\mu_0 - \left( z_{(1-\alpha/2)} \cdot \left( \frac{n}{\chi_{(1-\beta)}^2(n)} \right)^{\frac{1}{2}} \right) \cdot s_a \leq \mu_{a,y} \leq \mu_0 + \left( z_{(1-\alpha/2)} \cdot \left( \frac{n}{\chi_{(1-\beta)}^2(n)} \right)^{\frac{1}{2}} \right) \cdot s_a \right] \geq (1 - \alpha)\right\} = \beta \quad (24)$$

where  $s_a$  is estimated according to (16),  $\chi_{(1-\beta)}^2(n)$  means chi-quadrat distribution with  $n$  degree of freedom.

Based on the equation (24) we can say that for predefined values of measuring system's reliability  $(1 - \alpha)$  and dependability  $\beta$  and the number of measurements  $n$  the accuracy of measuring system will be:

$$\delta = \left( z_{(1-\alpha/2)} \cdot \left( \frac{n}{\chi_{(1-\beta)}^2(n)} \right)^{\frac{1}{2}} \right) \cdot s_a \quad (25)$$

### Unknown mean value and standard deviation

This variant is the most important in many practical cases, but the solution is theoretically very difficult. However, a lot of approximation forms exist based on which the practical simulation could be feasible.

We start by the task description

$$P\left\{P\left[\bar{\mu}_a - k \cdot s_a \leq \mu_{a,\gamma} \leq \bar{\mu}_a + k \cdot s_a\right] \geq (1 - \alpha)\right\} = \beta \quad (26)$$

where the sample mean value  $\bar{\mu}_a$  and sample standard deviation  $s_a$  are estimated from  $n$  samples according to (16).

Howe [4] defines a very simple approximation form for  $k$ :

$$k \approx \left(\frac{n+1}{n}\right)^{\frac{1}{2}} \cdot z_{(1-\alpha/2)} \cdot \left(\frac{n-1}{\chi_{(1-\beta)}^2(n-1)}\right)^{\frac{1}{2}} \quad (27)$$

Bowker [5] defines:

$$k \approx z_{(1-\alpha/2)} \cdot \left[1 + \frac{z_\beta}{\sqrt{2n}} + \frac{5 \cdot z_\beta^2 + 10}{12n}\right] \quad (28)$$

Ghosh [6] defines the next approximation form:

$$k \approx z_{(1-\alpha/2)} \cdot \left(\frac{n}{\chi_{(1-\beta)}^2(n-1)}\right)^{\frac{1}{2}} \quad (29)$$

If we take the approximation forms for  $z_x$  for  $x > 0.5$  [2]<sup>2</sup>:

$$z_x = u_x - \frac{2.30753 + 0.27061 \cdot u_x}{1 + 0.99229 \cdot u_x + 0.04481 \cdot u_x^2} \quad (30)$$

$$u_x = \left[\ln(1-x)^{-2}\right]^{\frac{1}{2}}$$

and for  $\chi_x^2(\gamma)$  [3]<sup>3</sup> (the number of degree of freedom is usually  $\gamma = n - 1$ ):

$$\chi_x^2(\gamma) = \gamma + z_x \cdot \sqrt{2} \cdot \gamma^{\frac{1}{2}} + \frac{2}{3} \cdot (z_x^2 - 1) + \frac{1}{9\sqrt{2}} (z_x^3 - 7 \cdot z_x) \cdot \gamma^{-\frac{1}{2}} - \frac{1}{405} (6 \cdot z_x^4 + 14 \cdot z_x^2 - 32) \cdot \gamma^{-1} + \frac{1}{4860\sqrt{2}} \cdot (9 \cdot z_x^5 + 256 \cdot z_x^3 - 433 \cdot z_x) \cdot \gamma^{-\frac{3}{2}} \quad (31)$$

or a much simpler approximation form from [3]:

$$\chi_x^2(\gamma) = \frac{1}{2} \cdot \left[ z_x + (2 \cdot \gamma - 1)^{\frac{1}{2}} \right]^2 \quad (32)$$

<sup>2</sup> The approximation error is not greater than 0.003

<sup>3</sup> For  $x \in (0.01, 0.99)$  and  $\gamma \geq 20$  the absolute error of approximation is not greater than 0.001

then the analytical equation for the estimation of measuring system's accuracy  $\delta$  based on an estimated mean value and standard deviation of  $n$ -sample data with the predefined measuring system's reliability  $(1-\alpha)$  and measuring system's dependability  $\beta$  can be computed.

#### 4.3 Illustrative example 1 - Simulation result

A very important question can be addressed with regard to accuracy: How the measuring system's accuracy  $\delta$  depends on the number of measurements for the prescribed measuring system's reliability and measuring system's dependability? We consider the following prescribed values for  $\alpha, \beta$ :

- a.  $\alpha = 0.3, \beta = 0.5$
- b.  $\alpha = 0.05, \beta = 0.99$

From (26) the measuring system's accuracy  $\delta$  is given  $\delta = k \cdot s_a$ . For finding the parameter  $k$  the equations (30), (31) and (32) were used. The Fig. 1 and Fig. 2 show the dependence of the parameter  $k$  on the number of measurements  $n$  for cases a) and b) respectively.

#### 4.4 Illustrative example 2 - Simulation result

Can the results from example 1 be proved by a simulation, if the number of measured values is  $n=30$ ?

For both cases a)  $\alpha = 0.3, \beta = 0.5$  and b)  $\alpha = 0.05, \beta = 0.99$  the values of  $k$  were found in Fig. 1 and Fig. 2 respectively:

- a.  $k=1.0584$ ,
- b.  $k=2.797$ .

In MATLAB, a set of 1000 samples of normal distribution with zero mean and standard deviations equal to one was generated. From the first 30 samples the mean value and standard deviation were estimated (16). Then the interval  $(\bar{\mu}_a - k \cdot s_a, \bar{\mu}_a + k \cdot s_a)$  was selected in accordance to (26) and the probability of falling into interval was computed from the whole set of 1000 samples - which is a procedure of how to compute measuring system's reliability.

The above mentioned procedure was repeated 5 000 times, and the probability of exceeding the predefined measuring system's reliability limit was computed - thus the test of measuring system's dependability was performed.

The obtained results for measuring system's dependability through simulation will be summarized as follows:

- a.  $\beta = 0.553$
- b.  $\beta = 0.9864$

The test was repeated many times and the predefined parameters for measuring system's reliability and dependability were achieved in all experiments.

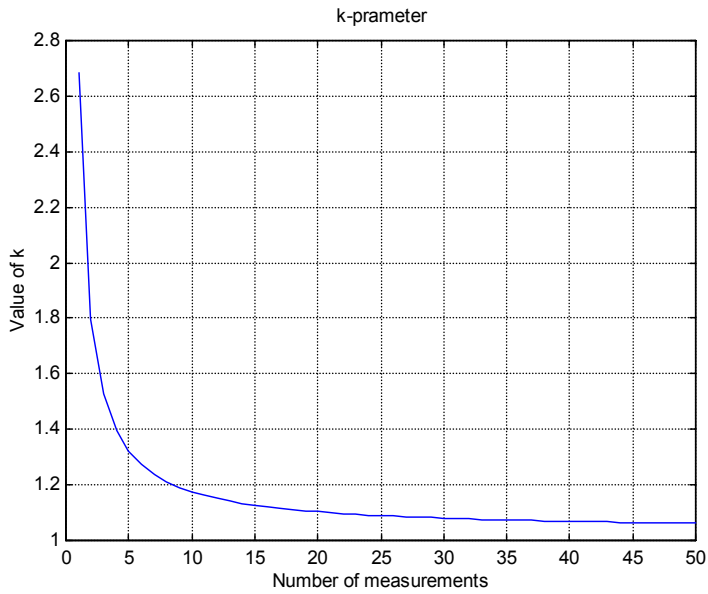


Fig. 1. Dependence of parameter k on number of measurements ( $\alpha = 0.3, \beta = 0.5$ )

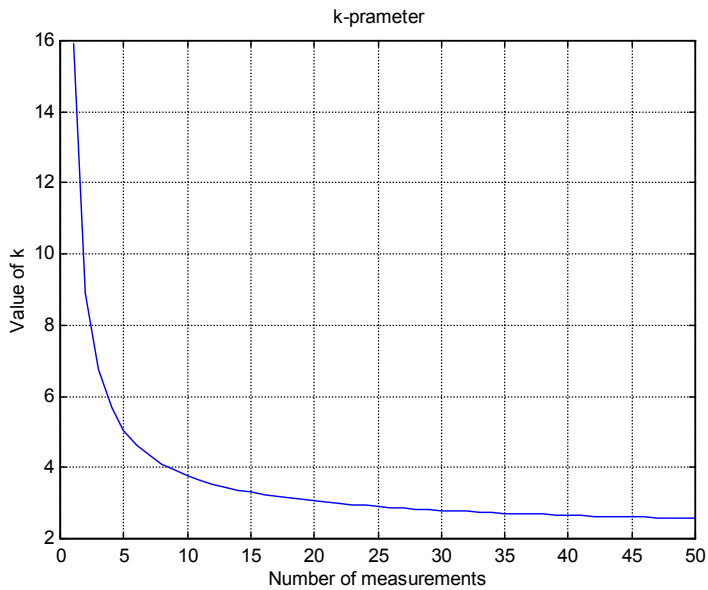


Fig. 2. Dependence of parameter k on number of measurements ( $\alpha = 0.05, \beta = 0.99$ )

## 5. Assessment of safety performance parameters

We can suppose  $N$  sensors data available where the probability of right error detection is marked as  $P_{RD}$  and the probability of non-correct error detection as  $P_{FD}$ . Because of the enormous safety and economical impact in case of non-correct error alert, the method of filtering "M from N" will be presented.

Let us have  $N$  sensors and for simplicity let us suppose the same probabilities of correct  $P_{RD}$  and non-correct error detection  $P_{FD}$  on each sensor. If this assumption is not fulfilled the method can be easily extended to a more general case.

As mentioned above, the hypothesis  $H_0$  represents perfect system behavior (non system error, no sensor error) and hypothesis  $H_1$  as a state with detected error (error of system, or error of sensors).

In the next equation, the probability of error detection on  $k$  sensors of  $N$  sensors ( $N-k$  sensors do not detect errors) is given in case the system does not display any error (conditioned by hypothesis  $H_0$ ):

$$P[k | H_0] = \binom{N}{k} \cdot P_{FD}^k \cdot (1 - P_{FD})^{N-k} \quad (33)$$

In the same way, the probability of error detection by  $k$  of  $N$  sensors is given in case the system is in an error state (conditioned by hypothesis  $H_1$ ):

$$P[k | H_1] = \binom{N}{k} \cdot P_{RD}^k \cdot (1 - P_{RD})^{N-k} \quad (34)$$

The main idea of "M from N" filtering is in selection of value  $M$  (threshold) defining the minimum number of sensors that detected error. If  $M$  sensors detect error then this error is taken as the real system error and the system starts sending error alert signals. The threshold  $M$  should be selected with respect to the following probabilities:

$$\begin{aligned} P_F &= \sum_{k=M}^N \binom{N}{k} \cdot P_{RD}^k \cdot (1 - P_{RD})^{N-k} \\ P_D &= \sum_{k=M}^N \binom{N}{k} \cdot P_{FD}^k \cdot (1 - P_{FD})^{N-k} \end{aligned} \quad (35)$$

where  $P_F, P_D$  means probability of a *false alert* (an error is detected but the system works without any errors) and the probability of the *right detection* (the system error is correctly detected).

The number of detectors  $N$  and the threshold  $M$  can be chosen based on sensors parameters  $P_{RD}, P_{FD}$  and required probabilities  $P_F, P_D$ .

Methods of data fusion and comparison are the main tools for estimation of system performance parameters (accuracy, reliability, integrity, continuity, etc.) and can be used for a derivation of an exact definition of false alert and right detection probabilities.



### 5.1 Illustrative example - Geo-object detection

In this example the measurement data comparison will be used as a tool for better geo-object detection in, e.g., electronic tolling application.

We can suppose  $N$  available position measurements of a geo-object where the probability of the right geo-object detection is marked as  $P_{RD}$  and the probability of non-correct (false) geo-object detection as  $P_{FD}$ . Let the hypothesis  $H_0$  represent the assumption of a perfect geo-object detection (no detection error reported). The hypothesis  $H_1$  represents a non-correct geo-object detection (error caused, for example, by wrong position accuracy, etc.).

The probability of  $k$  non-correct geo-object detections of  $N$  measurements for the final assumption that the geo-object is perfectly detected (conditioned on the hypothesis  $H_0$ ) can be given:

$$P[k | H_0] = \binom{N}{k} \cdot P_{FD}^k \cdot (1 - P_{FD})^{N-k} \quad (36)$$

The probability of  $k$  correct geo-object detections of  $N$  measurements for the final assumption of non-correct geo-object detection (conditioned on hypothesis  $H_1$ ) is given:

$$P[k | H_1] = \binom{N}{k} \cdot P_{RD}^k \cdot (1 - P_{RD})^{N-k} \quad (37)$$

The main idea of "M matches from N measurements" principle is in the selection of the threshold  $M$  with respect to the following probabilities:

$$\begin{aligned} P_F &= \sum_{k=M}^N \binom{N}{k} \cdot P_{FD}^k \cdot (1 - P_{FD})^{N-k} \\ P_D &= \sum_{k=M}^N \binom{N}{k} \cdot P_{RD}^k \cdot (1 - P_{RD})^{N-k} \end{aligned} \quad (38)$$

where  $P_F, P_D$  means the probability of a *false alert of geo-object detection* (the geo-object is detected even though the vehicle did not go through it) and the probability of a *right geo-object detection* (the right geo-object is detected based on the measured data, and the vehicle went through it).

The number of measurements  $N$  and the threshold  $M$  can be chosen based on the position probabilities  $P_{FD}, P_{RD}$  and the required probabilities  $P_F, P_D$ . Further discussion will be presented within an illustrative example below.

There are two parallel roads (one under tolling, the other one free of charge) and the distance  $D$  between them of 20 meters, as it is shown in Fig 3. The length  $L$  is supposed to be 1 kilometer.

In this example, we will try to tune the parameter  $M$  to increase the probability of the correct toll road detection in order to reach the expected value of more than 99%.

We expect a maximum vehicle speed of 200 km/h or 55 m/s. If the length is 1000 m and the GPS receiver monitors the position every second, we can obtain as many as 18 position measurements per one road. The road can be distinguished by GPS received with probability app. 70% (we can assign the measurement to the right road, if the error is lower than  $D/2$  which is in our case 10 meter - this accuracy is typically achieved by a GPS receiver at a probability level of 70%).

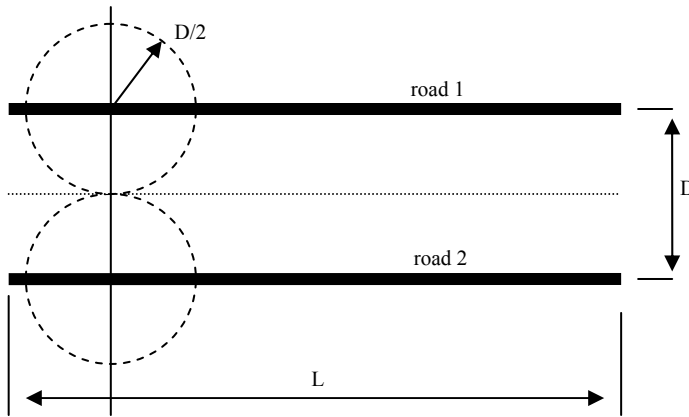


Fig. 3. Two parallel roads and Toll detection being on one of the roads

Based on the above mentioned assumptions, we can summarize the following parameters:

$$P_{RD} = 0.7, P_{FD} = 0.3, N = 18$$

Using the equations (38), the probabilities  $P_F, P_D$  for different parameters  $M$  will be as given in Tab.1.

Parameter $M$	$P_D$	$P_F$
6 from 18	0.9997	0.4656
8 from 18	0.9939	0.1407
10 from 18	0.9404	0.0210
12 from 18	0.7217	0.0014

Table 1. Probabilities  $P_F, P_D$  and their dependence on parameter  $M$

If the parameter  $M$  is 6 or 8, we can achieve the requested probability of the geo-object detection higher than 99%. On the other hand, for  $M=6$  the probability of lost vehicles is higher (the vehicles used the toll road, but the system did not detect them). For  $M=8$  we can achieve a better balance between both probabilities  $P_F, P_D$ . If the user needs to minimize the loss of a vehicle and to keep the acceptable detection probability, the variant  $M=10$  could be a good compromise.

## 6. Performance cluster of telematic applications using GNSS

Transport telematics architecture [22] displays the arrangement of subsystems and functional blocks, including information relationships according to the defined point of view. The task also covers the selection of representative telematics applications ("*cluster*") that shows identical systems requirements.

Among the individual representative applications using GNSS (Global Navigation Satellite Systems) the following may be included:

- Securing the movement of a means of transport in a transport infrastructure (from the point of view of performance parameters within the GNSS, it is a question of securing accuracy, reliability, availability, integrity, etc., at exactly defined points of the transport infrastructure - the application lays high stress both on the locator proper and the information transmission and processing systems; the solution should comply with the "fail-safe" principle; for typical transport telematics applications we may refer to railway interlocking technology, monitoring the transport of dangerous goods, or monitoring the movement of means of transport at the airport.
- Navigation of the means of transport in a transport network (from the point of view of performance parameters, it is a matter of coverage with a signal, time lag in on-line navigation, requirements as for the exactly working maps of an entire geographical area, requirements on the speed of information processing, both within a mobile unit and the processing centre, as well as minimisation of the delay when establishing the position - TTF - Time to Fix Face); as typical transport telematics applications, the following may be referred to: the navigation of safety and rescue units for a localised accident place or dynamic and/or on-line automobile navigation.
- Monitoring and operating the maintenance of transport networks (from the point of view of performance requirements, it is particularly a matter of an exact transport infrastructure information retrieval, interoperability of individual GIS (Geographical Information Systems) systems of various organisations dealing with maintenance, and achievement of high statistical accuracy in establishing position); as it concerns typical transport telematic applications, the following ones should be mentioned: mapping the river channel by means of a measuring ship, or measuring the carriageway parameters by means of special measuring vehicles.
- Monitoring the movement of persons and goods in a transport infrastructure (from the point of view of performance requirements, it is a matter of transmission and central processing of large amount of information from resources of various accuracy, fast identification of individual sub-sets of the objects of transport, sophisticated information processing in the centre, for instance, the "Floating Car Data"); as typical transport telematic applications, the following can be referred to: the use of taxi cabs, public transport passenger vehicles or other utility vehicles equipped with the GNSS systems for traffic flow modelling, or the use of localised mobile telephones for modelling the mobility of persons.
- Transport infrastructure charging according to its utilisation (from the point of view of performance parameters, it is a matter of reliability, integrity and time lag because the GNSS system is used to calculate the amount of the charge and, furthermore, the application places demands on the "fail-safe" principle in terms of the distance covered - if there is an uncertainty about correct charging of the driver, the distance covered is

not taken account of); as a typical transport telematic application, it is electronic charging of the transport infrastructure according to the vehicle parameters and distance covered.

As a follow-up to the completed analysis and decomposition of performance parameters to individual subsystems, a table can be obtained containing performance requirements of the above mentioned representatives as for the locator proper, telecommunications environment or the information processing centre.

### 6.1 Illustrative example - Assessment of telecommunications solution of GNSS monitoring system on airport surface

Described system methodology can be demonstrated on the telematic application based on GNSS (Global Navigation Satellite System) developed for the airport area moving objects management with good potential to be integrated into the already operated airport monitoring and management system.

The service central server collects and processes data received from all service vehicles. The obtained information is combined with the data gained from the existing systems. Processed and obtained result is distributed not the only to the airport management, but as well as to each vehicle equipped with active On-Board Unit (OBU) equipped with display of relevant size and quality. Each OBU receives also the managerial data generated by either airport control system or by dispatchers. Principal schema of the subsystem organization is displayed on Figure 4.

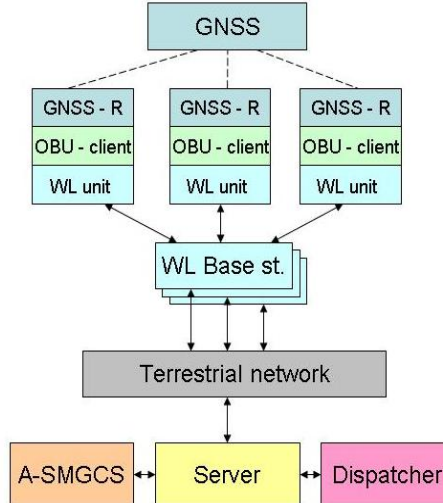


Fig. 4. Telematic service structure (WL – wireless, A-SMGCS - Advanced Surface Movement Guidance & Control System)

An airport area is precisely and transparently regulated area. The telematic sub-system performance indicators are introduced in Table 2.

Perform. Indicator	limit value	probability level	time interval
Accuracy	7.5m	99%	-
Availability	30s	99%	after init.
Reliability	36s	99%	3,600s
Continuity	5s	99%	180s
Integrity	5s	99%	-

Table 2. Required Telematics Performance Indicators in airport application

Using a transformation method described in [14] - [16] the telematic performance indicator "accuracy" was identified as the performance indicator with the dominant impact on the whole system performance. Its dominance is caused by the specific character of studied application. The requested level of accuracy (see Table 2) must be reached for every object moving with speed up to 120km/hour, If 1m GNSS sensor accuracy can be reached (the differential GNSS alternative must be applied) 195ms remain for the delay caused by all devices including the potential error healing in case of any sub-system problem (all on probability level 99% - see table 2).

Mobile WiMax (IEEE Std. 802.16d) was identified as the only possible alternative of the wireless access solution for the critical areas of the airport. All the other available access systems like GSM based products DTMF, HSCSD GPRS and EDGE as well as UMTS were identified as inappropriate. WiFi system operated in the open frequency band does not provide any service quality guarantee. Table 3 displays the obtained dynamical parameters of the WiMax channel (ART - Average Round Trip delay) in two critical stages of the Signal to Noise Ratio (SNR)

Site	Visibility	ART [ms]	SNR [db]
1	LOS	45.6	33
2	LOS	47.1	32
3	NLOS	44.6	-26
4	NLOS	44.8	-27

Table 3. Principle parameters of the WiMax access

Even though WiMax was selected as the core mobile access system for the airport critical areas, the whole airport area coverage with this technology is not economical. Some of alternative access solutions (EDGE/GPRS/UMTS or even WiFi) for Mobile WiMax difficult/irrelevant areas can be applied, if system parameters of these technologies meet these areas system parameters requirements. For such case multi-path solution like CALM, IEEE 802.21 based or by authors announced adaptive multi-path alternative (see e.g. in [20] - [21].) are identified as the appropriate alternatives.

The L2 ring based solution of the terrestrial chain part with the local QoS management (Hirschman) was applied to fulfill time limits of the whole chain.

The applied method nosily disqualified such communications “gurus” like the MPLS backbone terrestrial networking or GPRS/EDGE wireless access and combination of the wireless solution based on IEEE Std. 802.16d with terrestrial L2 switching solution with QoS management tools implemented were applied as core technology, they were tested and reasonable results were obtained. Critical issue, however, represent implementation of the effective decision processes to manage the multi-path solution to be kept in the required time limits – see [13] - [20].

## 7. Certification and testing of telematic applications

General methodology for telematic certification and testing of performance parameters [21] is figured on Fig. 5.

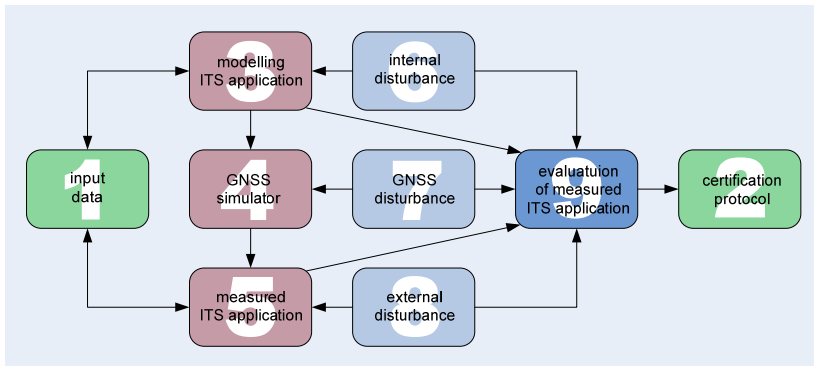


Fig. 5. General certification system for performance evaluation of telematic applications

Telematic application certification process consists of these steps:

1. definition of initial conditions for application (block 1);
2. optimally operating (modelling) telematic application definition (block 3);
3. disturbance statistics definition – internal disturbance of telematic application (block 6, disturbance of whole set of tested vehicles or OBUs), disturbance of GNSS signal (block 7) and external disturbance (block 8);
4. activation of measured (certified) telematic application (block 5);
5. testing of measured telematic application with simulated GNSS signal (block 4) for all defined situations/scenarios (initial conditions – block 1) with goal to cover all suitable situations;
6. real testing of selected (available) scenarios;
7. conformity assessment of output data from appropriate (tested) telematic application and output data from model (optimally operating) application for defined initial conditions and defined disturbances;
8. measurement results processing for certification protocol (protocol of measurement) and final assessment of telematic system parameters guarantee (block 2);
9. performance parameters guarantee must be statistically verified on a sufficient number of measurements to be able to guarantee monitored properties in defined statistical parameters.

### 7.1 Illustrative example - Pilot certification of telematic applications using GNSS

Pilot tests of GNSS signal reception conditions and their evaluation were carried out during September and November 2009 at e-Ident laboratory in Prague [21], the laboratory for electronic identification systems and communications. The signal was simulated with Spirent GNSS Simulator, type GSS 8000. For pilot testing the GPS positioning system was selected as a reference GNSS system.

The routes were generated by creating NMEA messages (a special file was created for each route), every case is located in Prague area and its nearby neighbourhood. NMEA log files of routes for pilot testing were recorded during September in real test rides in cars, with the GPS unit. This data was processed and adjusted according to requirements for pilot testing in laboratory.

For virtual gate passage tests the hardware equipment of universal telematic mobile unit as OBU (On-Board Unit) was used. Installed DEFT application (Dynavix, EFC, Fleet, Toll) was also developed on testing hardware device.

For testing were chosen 2 following testing sections. Both sections are situated in the west of Prague, where both directions of travel through the section are considered and prepared for testing (Fig.6):

- section of road II/605 around the bridge over the highway R1; to initialize the GNSS unit, previous section of road II/605 (in the direction from the city centre) was used; section transit speed between 50 and 70 kph;
- section of highway R5 near the Metropole shopping centre; to initialize the GNSS unit, previous sections of highways R1 and R5 were used; section transit speed between 80 and 100 kph;

Test sections have been recorded into the log file in GPS unit placed on the windscreen of testing vehicle. This data was then processed into various NMEA files (each represents one testing section and one direction of travel).

The NMEA records were cut and connected into series. Series of measurement consist of 10 experiments, 5 of them in one direction and 5 of them in the opposite direction, so that it can test the entire sequence, i. e. all 10 passages through the defined segment. Before each series testing, there was simulated passing of the initialization section (to get fix of GPS signal and calibrate the equipment).

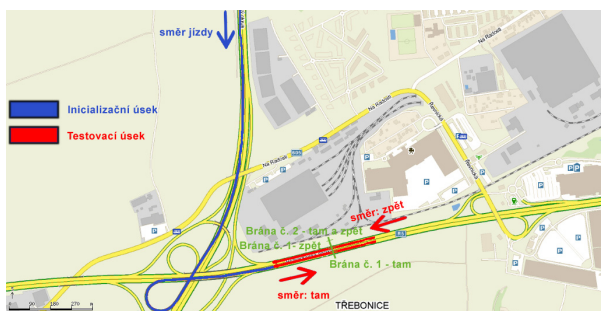


Fig. 6. Initialization (blue) and testing (red) section on highway R5, western part of Prague

Virtual gates locations were defined by geographical coordinates and the direction was determined by the starting azimuth of drive. For testing there was set azimuth angle of  $\pm 90^\circ$  (from the road axis,  $45^\circ$  on each side).

Virtual gate passage was detected with software by the principle of measuring the distance from defined positions of the virtual gate in the outer circle of radius  $r_1$ , which takes the value of the distance travelled by a vehicle at time  $t_1 = 2$  s, respectively, crossing the border of inside circle of radius  $r_2$ , which takes the value of the distance that the vehicle travels at time  $t_2 = 1$  s (Fig.7).

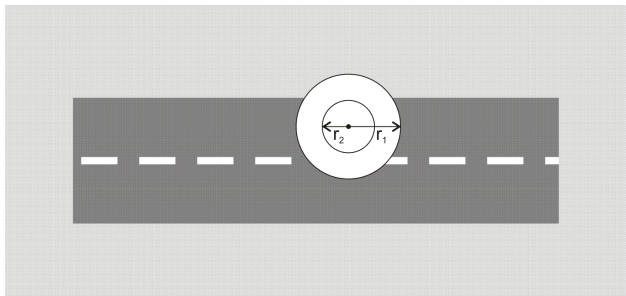


Fig. 7. Radius of detection circles

The geographical coordinates defining the location of the virtual gates were deduced from the publicly available maps on Internet, the azimuth was derived from the test drive data. Number of defined virtual gates was not restricted. For each section have always been defined 2 positions of virtual gate in each direction:

- the first on the passing road;
- the second in order of tens of meters from the passing road for the simulation of parallel road, where it is possible to expect the most common mistakes, so it is a gate on a fictitious road (for parallel road testing does not matter whether the road is fictitious or real).

For the section on road II/605 the location of the second gate (parallel communication) was defined at a distance of approx. 20 m from the axis of the road. In the case of section on highway R5 the virtual gate for parallel communication is located at a distance of approx. 30 m from the axis of one belt of the highway.

The virtual gate passage detection testing (with usage of GNSS simulator) was carried out at a total of 26 series of 10 measurements (Fig.8):

- 15 series for the section of road II/605, i. e. a total of 150 measurements,
- 11 series for the section of highway R5, i. e. a total of 110 measurements.

The tests were carried out at first with the default GNSS signal without any restrictions and with the defined parameters of the route. After that further test scenarios were developed and the simulated signal was influenced in the following way:

- changing in signal power from different satellites;
- turning off selected satellites;



- simulation of built scenes of surroundings – highway, city, suburbs – which contain a typical set of ground clutter, multipath signal transmission and signal shading;
- simulation of various predefined atmospheric changes.

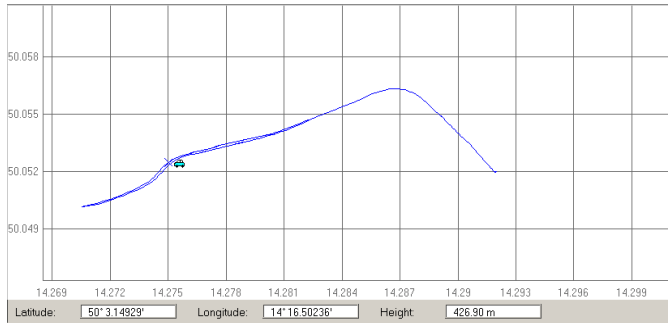


Fig. 8. Simulation output of the position on road II/605

The simulated signal was influenced at first for whole measurement series, this approach was later changed to influencing the signal for various tests in the series, always another way, such as switching on and off the satellites at a defined angle above the vehicle, respectively above the ground (road plane). Signal influence has been prepared based on detailed description of possible signal influence. For the tests the spherical characteristics of the receiving antenna were chosen, with an open top of the antenna (i. e. position of the antenna on the roof of the vehicle).

Running of the tests has been recorded in the “log file” stored directly in the hardware unit. Based on this data further processing took place already on PC. The hardware unit recorded the following data of a virtual gate passage:

- virtual gate passage time;
- ID and a description of the gate (defined positions and azimuth).

Each measurement result was then classified as “passed” or “failed” according to the following categorization:

- **passed**, if all the following conditions are true:
  - gate is identified on the running road;
  - gate is not identified in other nearby road;
  - any upstream gate is not identified;
  - more passages through the same gate during one test are not evaluated;
- **failed** in other cases.

Results of carried out testing are listed in Tab.4 [22]. Based on the measurements, it was demonstrated, that the results of the passage identification varies for different parameters of the GNSS signal, environment and other influences on the signal reception. The relatively low percentage of successful running gate identification may be due to high sensitivity software in the OBU. It can be assumed that for usage of OBU for telematic applications the higher success rate for negative detection of upstream or neighbouring gates will be demanded.

Section and direction	Type of test	Passed in %	Failed in %
Road II/605 forward	Passed gate	60.0	40.0
	Upstream gate	100.0	0.0
	Nearby downstream gate	97.3	2.7
	Nearby upstream gate	98.7	1.3
Road II/605 backwards	Passed gate	50.7	49.3
	Upstream gate	98.7	1.3
	Nearby downstream gate	96.0	4.0
	Nearby upstream gate	100.0	0.0
Highway R5 forward	Passed gate	65.5	34.5
	Upstream gate	96.4	3.6
	Nearby downstream gate	100.0	0.0
	Nearby upstream gate	100.0	0.0
Highway R5 backwards	Passed gate	69.1	30.9
	Upstream gate	100.0	0.0
	Nearby downstream gate	100.0	0.0
	Nearby upstream gate	100.0	0.0
On the whole	Passed gate	60.4	39.6
	Upstream gate	98.8	1.2
	Nearby downstream gate	98.1	1.9
	Nearby upstream gate	99.6	0.4

Table 4. Results of pilot testing, depending on the segment, direction and type of test

Within implementation of a sufficiently large count of measurements the resulting values in Tab.4 show the probability of conformity of the tested OBU properties with the desired properties of the measured applications.

Proposed test protocol is divided into a part of recording the individual measurements and a part of evaluation of the application as a whole, including evaluation of system parameters. It is already possible on the basis of pilot testing to summarize the partial results and define partial requirements for selected system parameters of tested applications:

- a. **Accuracy** - the required value of 15 meters in the horizontal plane at 95 % level of probability. Thus defined accuracy corresponds to usual accuracy requirements on OBU GPS unit using a standard statistical distribution of signal parameters. The tested OBU and its software configuration were selected based on the required accuracy. Accuracy of the OBU can be further increased.
- b. **Reliability** - the specific desired value will be defined on basis of experience in the follow-up testing. Assuming a value of around 90 %, eventually higher, at 95 % probability level. After pilot tests result with 60.4 % success of virtual gate passage detection can be seen that the higher reliability will be required for the approval. To approve the usage of tested OBU it will be necessary to adjust the software in OBU or to use another OBU. Reliability is affected by the security parameter, see below.
- c. **Availability** - the specific desired value will again be defined on the basis of experience in further testing. Assuming the OBU unit activation at start-up of travel and the availability value up to 60 seconds. Pilot testing showed in most cases the availability of tens of seconds.
- d. **Continuity** - for telematics application this is not a critical system parameter. Depends on ability of the telematic system to assess and calculate the travelled route in case of system failure while driving the vehicle; this parameter does not depend only on the OBU and GNSS signal reception. In case of the usage of virtual sections or more virtual gates on the road sections the requirements on this parameter are significantly decreasing.
- e. **Integrity** - this parameter has not been considered for the tested systems, mainly because it represents rather the quality of OBU diagnostics, resp. informing the user within a reasonable time for failure of proper OBU function. This parameter is not so much related to position determining or frequency of the virtual gates and sections passage records.
- f. **Safety** - "dangerous conditions" are defined for both tested applications in the field of electronic toll system (EFC). This is a situation where the system assesses charges, which in fact did not occur. This is part of the conditions to test status "failed" - so-called "false alarms" - i.e. the identification of neighbouring gates or division, identification or evaluation of multiple transits the same gate or the same road section. With regard to safety (as one of the important system parameters of tested applications) appropriate testing scenarios were selected by pilot tests - all the scenes are containing unpleasant situation in a similar gate on nearby parallel road (i.e. identically oriented and situated at a distance of tens of meters from the passing road). It should be emphasized that effort to eliminate false alarms is associated with a reduction on the reliability parameter of applications, i.e. reducing the probability of passed gates/sections correct detection. It is therefore necessary to seek to balance, which on the one hand significantly reduces the risk of false positives, on the other hand, provides useful reliability parameter. This balanced condition can be found by testing various OBU units, resp. by testing one OBU at various software settings.

## 8. References

- [1] Test of normality  
[[http://www.caspur.it/risorse/softappl/doc/sas\\_docs/qc/chap1/sect19.htm](http://www.caspur.it/risorse/softappl/doc/sas_docs/qc/chap1/sect19.htm)]
- [2] Novovicova J.: *Probability and mathematical statistics*, textbook, Faculty of transportation sciences, CTU Prague, 1999.
- [3] Jilek M.: *Statistical confidence intervals*, Teoreticka kniznice inzenyra, SNTL, Prague, 1988 [in Czech].
- [4] Howe W.G.: *Two-sided tolerance limits for normal populations*, JASA, 64, 1969.
- [5] Bowker A.H.: *Computation of factors for tolerance limits on a normal distribution when sample is large*, AMS 17, 1946.
- [6] Ghosh D.T.: *A note on computation of factor for tolerance limits for a normal distribution*, Sankhya B42, 1980.
- [7] COST 323, *European Specification on Weigh-In-Motion of Road Vehicles*. Draft 3.0, EUCO-COST/323/x/99, June 1999.
- [8] Jacob, B., *Assessment of the Accuracy and Classification of Weigh-In-Motion Systems*, Part 1 - Statistical Background. March 1997.
- [9] Lieshout, R.A. van, Zoutenbier, M.H.H., *Weigh-In-Motion of Road Vehicles*. WIM-VID/IMP1, nr 6800.0770, E1655-01, CQM. May 1998.
- [10] Lieshout, R.A. van, Zoutenbier, M.H.H., *Weigh-In-Motion of Road Vehicles*. WIM-VID/IMP1, nr 68000.0821- Studying Measurement Accuracy-, E1657-01, CQM. Sept 1999.
- [11] Owen, D.B., *Handbook of Statistical Tables*. Reading, Mass, 1962.
- [12] Stig Danielson: *Accuracy in WIM systems - An examination of different methods for determining precision*, Report, Linkoping University, Department of Mathematics Statistics
- [13] M. Svitek,, T. Zelinka, T., *Communications Solutions for ITS Telematic Subsystems*, WSEAS Transactions on Business and Economics, Issue 4 (2006), Vol. 3, pp 361 - 367, Athens 2006, ISSN 1109-9526,
- [14] T. Zelinka, M. Svitek, *Communication solution for Vehicles Navigation on the Airport territory*, Proceedings of the 2007 IEEE Intelligent Vehicle Symposium, Istanbul, Turkey, pp 528-534, IEEE Catalogue number 07TH8947, ISBN 1-4244-1068-1.
- [15] T. Zelinka, M. Svitek, *Communication Scheme of Airport Over-ground Traffic Navigation System*. ISCIT 2007. IEEE Sydney, 2007, pp 329 - 334. IEEE Catalogue No. 07EX1682(C), ISBN 1-4244-977-2, Library of Congress 2007920360.
- [16] M. Svitek,, T. Zelinka, *Monitoring of Transport Means on Airport Surface*. Advances in Transport Systems Telematics, edited by Jerzy Mikulski, Selesian University of Technology, Katowice, pp. 285 - 292, ISBN 978-83-917156-6-6.
- [17] T. Zelinka, M. Svitek, *Decision processes in telematic multi-path communications access systems*. International Journal of Communications, North Atlantic University Network NOUN, Issue 2, Volume 1, 2007, pp.11 - 16.
- [18] M. Svitek,, T. Zelinka, *Communications multi-path access decision scheme*. Neural Network World, ICS AS CR and CTU, FTS, Praha, No. 6.,2008, pp 3 - 14, 2008, ISSN 1210 0552,
- [19] T. Zelinka, M. Svitek, *Adaptive communications solutions in complex transport telematics systems*. Computers and Simulation in Modern Science- II, WSEAS Press, Athens 2009, pp. 234 -241, ISBN 978-960-474-032-1.

- 
- [20] T. Zelinka, M. Svitek, M. Srotyr, M. Vosatka: Adaptive Multi-path Telecommunications Solutions for ITS, *Journal of Systemics, Cybernetics and Informatics* Volume 9, No. 1, pp. 14 - 20, Orlando, 2011, ISSN: 1690-4524.
- [21] Svitek, M; Faltus, V; Lokaj, Z: Pilot Testing of Certification Methods Developed for ITS Applications Based on GNSS Systems, 13th International IEEE Conference on Intelligent Transportation Systems, Madeira, PT, 2010, ISBN 978-1-4244-7658-9, ISSN 2153-0009.
- [22] M. Svitek: Intelligent Transport Systems - Architecture, Design methodology and Practical Implementation , [Key-note lesson], World Scientific and Engineering Academy and Society, 5th WSEAS/IASME Int. Conf. on Systems Theory and Scientific Computation, Malta, 2005.

## **Section 2**

### **Length, Distance and Surface**



# Length Metrology and Calibration Systems

Saulius Kaušinis<sup>1</sup>, Albinas Kasparaitis<sup>2</sup> and Aurimas Jakštas<sup>1</sup>

<sup>1</sup>*Kaunas University of Technology*

<sup>2</sup>*Vilnius Gediminas Technical University  
Lithuania*

## 1. Introduction

Accurate length measurement plays a vital role in meeting the needs of industry and commerce for traceability to common national and international standards, especially in view of the common market and world trade. Such measurement needs arise across a wide applications base, from large-scale engineering projects such as dam construction, aerospace and shipbuilding, through automotive engineering and components manufacture, to precision engineering and nanotechnology (DTI/NMDS, 2002b).

The lowest uncertainty attained in dimensional measurements of a material object occurs in semiconductor industry and integrated circuit (IC) production. The dimensional feature of interest in a line scale is the critical dimension (CD). The CD corresponds to the width of the smallest line that can be produced on a wafer with an acceptable yield of manufactured devices; presently this parameter is less than 0.1  $\mu\text{m}$ . Requirements in other areas, such as manufacture of precision instruments, large machines (e.g. planes), and others also rise fast. In all these areas the principle “to stop means to fall behind” is in force. Development of measurement systems is impelled by the augmentation of customer needs as well as by steadily evolving state-of-the-art measurement technologies (Bosse & Flügge, 2001).

Length metrology has a fundamental role to maintain the primary standard of length, the metre, and to provide the infrastructure to enable a wide range of dimensional and positional measurements to be made traceable to the metre. National metrology institutes (NMIs) in a number of countries and companies that produce precision high-tech products pay much attention to accuracy-related research with the aim to improve properties of length calibration systems and to specify their uncertainty budget. Metrological programmes in the area of length measurement are consistently carried out in the USA, Japan, UK, Germany (Bosse & Flügge, 2001; Beers & Penzes, 1999; Israel et al., 2003), and other countries. The programmes impel the creation of metrological infrastructure that increases industry competitiveness, supports industrial innovations, and improves control of manufacturing processes and quality. For example, systematic research of accuracy of vacuum nano-comparator, performed in German National Metrology Institute (PTB) in 2000 – 2006, resulted in reducing the measurement repeatability error from 14 nm down to 0.2 nm. NIST, the National Metrology Institute of the USA, is carrying out research on nm-accuracy one dimensional (1D) metrology with the development of components of next generation length scale interferometer. In conceptual design, the system would have a range for 1D measurements from 100 nm to 1 m with a target expanded uncertainty of from 1 nm to 10 nm.



One of the most sophisticated challenges for science and the high technologies engineering is the growing need to address real industrial problems rather than the ideal measurement situation and embed the traceable length metrology directly into technological processes by performing precise dynamic measurements in more demanding environments than those of calibration laboratories.

This chapter will present a synopsis and analysis of literature and existing scientific and technical solutions of precision length calibration. It covers analysis of laser interferometers, line detection systems, measurement signals and algorithms, as well as measurement capabilities of state-of-the art calibration systems worldwide. The contribution also addresses a thorny issue of achieving reliable measurements and meeting contradictory requirements of accuracy and productivity of line scale calibration in non-ideal environmental conditions, under the influence of many external influencing factors. The problems will be also upon the anvil of the development of an interferometer-controlled comparator that is operated in dynamic mode and enables to trace the calibration of line scale of up to  $L \leq 3.5$  m long to the wavelength standard.

## 2. Length metrology

The basis of any dimensional measurement technique is found on the realization of the SI unit of length via frequency-stabilized lasers and displacement interferometry. The measurement technologies employed include laser-ranging devices, large-scale coordinate measuring machines (CMMs), optical- and ultraviolet-light microscopes, scanning electron microscopes (SEMs), atomic force microscopes (AFMs), and scanning tunneling microscopes (STMs).

Both direct and indirect high accuracy measurements of length, distance and displacement make use of wavelength or optical frequency sensing techniques. Direct measurement techniques include laser interferometer calibration of computer numerical controlled (CNC) machine tools and CMMs, and commercial laser-based instrumentation is widely used both nationally and internationally for this purpose, to measure displacements and distance from typically a hundred nanometers to tens of meters. Multiple wavelength instrumentation is used to extend accuracy within well-controlled environments, whilst modulated laser ranging techniques (electronic distance measurements) are now widely applied in surveying over distances up to a few kilometers with, in some cases, sub-millimeter precision. Such precision instrumentation comprises laser wavelength sources as measurement transducers of varying degrees of stability and accuracy.

Dimensional metrology covers measurement of dimensions and in principle also geometries based on distance measurements in a wide range of more specific measurements, targeted on from primary sources, i.e. lasers to geometrical measurement of complex profiles, which typically include:

- measurement of laser wavelength/frequency, stability, drift and line width of radiation sources that are used for interferometry and distance measurement;
- measurement of size or geometric features, like pitch, of 1D artifacts, for example end standards and linear scales or encoders;
- measurement of size and/or locations of features in 2D structures common in the semiconductor industry, such as in the complex patterns of integrated circuits
- measurement of size location and orientation of features in 3D patterns;

- measurement of deviations from ideal geometric forms, i.e. flatness, roundness, etc.;
- measurement of surface texture.

Calibration of a variety of parameters associated with the source, such as absolute wavelength or frequency, linewidth, stability or drift, are thus of primary importance to high precision length traceability. In parallel, techniques in wavelength metrology are targeted on other applications. These include spectral bandwidth characterization by wavelength division multiplexing (WDM) for optical communications, high resolution spectral analysis using Fabry-Perot standards, and high accuracy measurement of spectroscopic phenomena, which has strong input to scientific spectroscopy. Precision length metrology also plays a key role in the realization of derived units of pressure and current, for example. The highest wavelength/frequency accuracy and stability available contributes to the leading-edge determination of fundamental physical constants (DTI/NMDS, 2002a).

There are a number of sensor technologies and instruments with nanometer, or better, accuracy for measuring length that repeat well if used carefully, including the scanning probe and electron microscopes and some optical devices. However, universal measurement standards have not yet been established and even apparently sophisticated users of atomic force microscopes can produce large variations in their measurements of the same artifacts. Without agreed standards, tools or machines cannot be calibrated at the nanometer scale (Bureau International des Poids et Mesures, 2003).

Line graduated geometric bodies, with graduation spacings representing known distances, are the bases for all direct measurements of specific distances. It follows that instruments having line graduated elements as integral members may be considered the only mechanical means capable of carrying out direct measurements without complementary equipment or processes (Farago, F.T. & Curtis, 1994).

The need for reduced uncertainty in the “primary standard” aspect of length, i.e., in its definition and realization, and in the “secondary standard” aspect, i.e., in its transfer and dissemination through dimensional metrology, is linked strongly to tightening tolerances in industrial manufacturing.

## 2.1 Definition and realization of meter

The definition of the meter—whether in terms of a prototype meter bar, a wavelength of light, or the propagation of an electromagnetic wave in an interval of time—has provided the basis for the lowest-uncertainty realization of the unit.

In 1983, the meter was re-defined again to the one in effect today, namely: “The meter is the length of path traveled by light in vacuum during the interval of  $1/299\,792\,458$  of a second”. At that time, the International Committee on Weights and Measures (CIPM) gave three basic methods for the practical realization of the meter: time-of-flight, using time intervals, and interferometry, using wavelengths or frequencies. CIPM gave five recommended radiations with assigned frequencies, wavelengths, and uncertainties (Quinn, 2003).

Of the recommended radiations, that of the iodine stabilized helium-neon (He-Ne) laser is the most widely used for practical realization of the meter. It has a wavelength of He-Ne = 632.991 398 22 nm, with a relative standard uncertainty  $u_r$  of  $2.5 \times 10^{-11}$ .

The effect of the re-definitions and advances in measurement of the frequencies of recommended radiations was to decrease the relative uncertainty attainable in realization of the meter by five orders of magnitude (Swyt, 2001).

Measurements of dimensions of material goods are most often referenced to the SI unit of length through material artifacts calibrated as dimensional standards. The meter, the basic unit for length, is usually transferred to measurement standards in the form of line scales or photoelectrical incremental encoders by length measuring machines that typically use a laser interferometer in air as reference measuring system. The measurement results are traceable to the meter due to the use of the wavelength of the laser interferometer.

## 2.2 Laser interferometry

Since practical realization of meter is closely related with the radiation of stabilized frequency lasers, laser interferometers are utilized for precise and traceable length measurements. Currently the detection principles of laser interferometer systems can be distinguished between homodyne and heterodyne techniques (Webb & Jones, 2004; Chapman, 2002). Homodyne interferometers utilize one frequency laser, and heterodyne two frequencies laser respectively. Heterodyne interferometry is inherently more resistant to noise due to its heterodyne frequency and the design of common-mode rejection which cancels out common noises coming from both reference and measurement signals, and though heterodyne techniques are susceptible to larger nonlinearity errors a large number of commercial systems uses namely this technique. Nevertheless the main parameters that determine the quality of laser interferometric systems are (Bobroff, 1993):

- resolution,
- measurement accuracy,
- repeatability of results,
- dynamic and measurement range,
- measurement speed.

A homodyne laser source is typically a He-Ne laser with a single frequency beam as output consisting of either a single polarization under  $45^\circ$  or a circularly polarized beam. The beam is split into the reference arm and measurement arm of the interferometer by a beam splitter. Following a reflection off their respective targets, the beams recombine in the beam splitter. In order to observe interference the two beams must have equal polarizations. This is accomplished using a linear polarizer oriented at  $45^\circ$  to the beam splitter. The photo detector signal is run through electronics which count the fringes of the interference signal. Every fringe corresponds to a path difference of half a wavelength. After superposition of measurement and reference beams a polarizing beam splitter is used to generate two  $90^\circ$  phase shifted signals. The direction of movement is determined at zero crossings of the interference signal using the other signal. Counting of the zero crossings of both interference signals provides a resolution of  $\lambda / 8$  which is not sufficient for precision length measurements and therefore it has to be enhanced by interpolation techniques. In homodyne interferometers the amplitudes of the interference signals are used; the phase of the signal can be determined from intensities of perpendicular polarized signals. Manufacturers of homodyne interferometers are Renishaw, Heidenhain, Sios and recently Interferomet.

In heterodyne interferometers double frequency radiation source is required since the interfering measuring and reference beams must have slightly different frequencies and photo detectors detect the phase shift between these two beams, see Fig. 1.

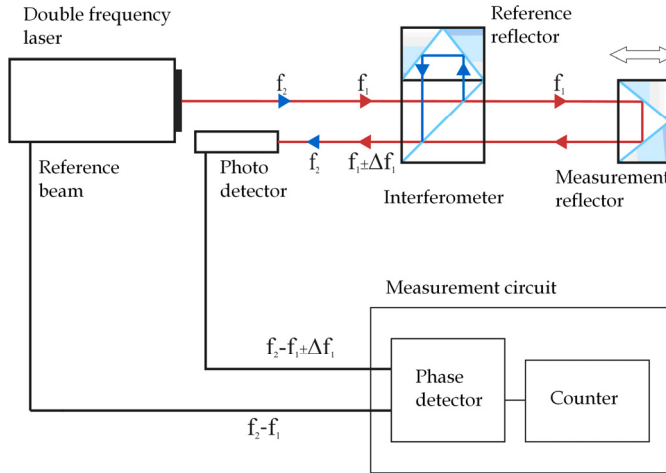


Fig. 1. Layout of heterodyne laser interferometer

Two frequencies are separated by their polarization state, so that a polarization beam splitter can generate a measurement beam with the frequency  $f_1$  and a reference beam with  $f_2$ . The movement of the measurement reflector with velocity  $v$  causes a frequency shift  $\Delta f$  in the measurement beam due to Doppler effect. This shift will increase or decrease depending on the movement direction of the measurement reflector. When counting the periods of reference and measurement signals simultaneously their difference is proportional to displacement. The direction of movement can be determined directly from the sign of this difference. Interpolation in heterodyne systems is equivalent to the measurement of the actual phase difference between fixed reference frequency and measurement frequency (Webb & Jones, 2004).

Heterodyne principle has a certain advantage, particularly for He-Ne lasers due to good signal to noise ratio enabling multi-axis measurement with high measurement speeds. In practice, the maximum speed of the laser interferometer will be limited by the primary beat frequency of a laser source, which is usually in the range of few MHz for Zeeman stabilization technique where high power magnetic field is required for separation of the laser frequencies, and 20 MHz for the acousto-optical modulators and from 600 to 1000 MHz for stabilized two-mode lasers that would correspond to more than 200 m/s speed to be measured. However, increase of the beat frequency will decrease the measurement resolution of the laser interferometer in return and therefore the bandwidth of phase detection is usually fixed in the absolute time scale in most phase measuring techniques (Yim et al., 2000).

The main advantage of heterodyne systems is that information about the measured displacement is obtained in form of variable signal and therefore measurement circuits are

not sensitive to variations of the measured signal level due to various disturbances. Since information about the displacement is gained from the signal frequency only one photo detector is necessary, and adjustment of optical elements becomes more simple, see comparison of both systems in Table 1.

	Homodyne	Heterodyne
Continuous measurement	Yes	No
Sensing of moving direction	Quadrature	Always
Quadrature output signal	Yes	Yes
Error detection	not defined	Unambiguous
Sensitivity to radiation intensity	Yes	No
Sensitivity to environmental irradiance	Yes	No
Frequency band of electronic circuit	$0 - 2v/\lambda$	$f_1 - f_2 \pm 2v/\lambda$
Signal to noise ratio	6 - 12 bit	2 - 3 bit
Multi-axis measurements	Limited	Yes
Photo detector	Complex	Simple
Adjustment simplicity	No	Yes

Table 1 Comparison of heterodyne and homodyne systems

Resolution of interferometers primarily depends on accuracy of phase detection of interference signal. Currently due to rapid development of information technologies and electronics it increased from  $\lambda/8$  in 1965 up to  $\lambda/2048$  (the wavelength of radiation source is divided into 2048 parts), and using traditional phase detection techniques the resolution of displacement interferometry amounts to about 0.1 - 10 nm.

Presently some commercially available phase detection systems allow digital signal processing and phase detection accuracy of  $0.01^\circ$  and which in combination with Michelson type heterodyne interferometer correspond to better than 10 pm system measurement accuracy.

The uncertainty sources in the laser interferometry can be grouped in three categories: setup dependent (cosine, Abbe, deadpath errors, mechanical stability), instrument dependent (stability of laser frequency, electronics, periodic deviations, etc.) and environment dependent (refractive index, turbulence, thermal sources) that are inherent in such systems and generally limit the relative uncertainty to  $2 \times 10^{-8}$ , resulting in an error of 20 nm per meter. For measurements over large displacements in air the last group is predominant in the uncertainty budget.

Thus linear measurement scale and accurate and adequate control and stabilization of environmental conditions are the main precision criteria for the modern interferometric displacement measurement systems. The use of digital data processing techniques enables to minimize nonlinearities of laser interferometers and increase the measurement resolution up to 10 pm (Webb & Jones, 2004). But in order to achieve the required relative length measurement uncertainty below  $10^{-7}$  in practical applications, the measurement accuracy of air refractive index must be not less than  $10^{-8}$ .

A good example for that is the comparison of static stability of He-Ne laser interferometer and linear encoder presented in Fig. 2. Experiments conducted have shown that users can expect fewer measurement fluctuations of the position display from linear encoders than from laser interferometers (Kaušinis et al., 2004).

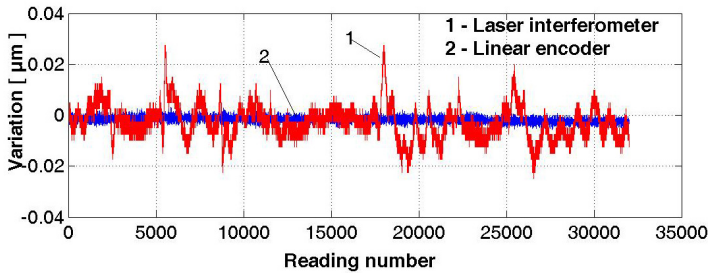


Fig. 2. Stability of laser interferometer and reference encoder

A long-term stability of the laser interferometer was mainly influenced by the temperature and refractive index. The refractive index of air was compensated by Edlén's formula. Both measurement systems were simultaneously read out in a static mode and actually, in comparatively good ambient conditions the interferometer has revealed clearly higher variances than the linear encoder. The temperature measurement system due to its data acquisition time and remoteness of the laser beam sensor is not able to compensate these variations.

The analysis of the dynamic mode of operation displayed results of the same order as variations of the laser interferometer measured in a static mode. Repeated measurements at different positions of the line scale displayed quite similar results. It is evident that, even in the finely air-conditioned laboratory environment, the low heat capacity of the air causes quick changes in temperature that can lead to relatively large fluctuations in measured values obtained from the laser interferometer. Comparatively short distances between encoder's scale and index grating minimize sensitivity to environmental factors.

### 2.3 Limits and challenges in length metrology

The future of length and dimensional metrology is being shaped by theoretical and practical limits to attainable uncertainties in measurement, by continuing trends in industry. There are two drivers that force the achievement of ever-smaller uncertainties in length and dimensional measurements. These are, first, the continuing industrial trend to tighter tolerances—represented in the microelectronics domain by Moore's Law—and, second, the continuing scientific trend to explore the limits of understanding through physical measurement.

The latter is bounded by a dimensional equivalent of Johnson, or thermal, noise that places an ultimate limit on the uncertainty of measurement of dimensional features. Thermal length fluctuations of a solid artifact, the spatial equivalent of electronic Johnson noise, are due to thermal agitation of the atoms of the material. In a measuring machine, such thermal noise places an ultimate limit on the ability to set the location of the origin of the axes of the

machine and, therefore, on the uncertainty of position measurements the machine can attain. Thermal noise similarly limits the uncertainty with which the length of an object can be measured. Therefore thermal fluctuation  $\Delta l$  in the length  $l$  of the side of the artifact is given:

$$\Delta l = \left( \frac{k \cdot T}{3} \cdot B \cdot l \right)^{1/2} \quad (1)$$

where  $k$  is the Boltzmann constant and  $T$  is the thermodynamic temperature,  $B$  is the bulk modulus of the material of the cube.

For example, for an object with a bulk modulus of that of fused silica,  $3.5 \times 10^{10}$  N/m<sup>2</sup>, and a temperature of 300 K, the rms fluctuation in dimension of a 1 m cube is 0.2 fm ( $10^{-15}$  m) or, fractionally,  $2 \times 10^{-16}$ . Besides this factor is inversely proportional to the measured length and therefore by decreasing the geometrical dimensions of measured objects the length measurement uncertainty will increase proportionally (Swyt, 2001).

Generally the limit for length measurement uncertainty is firstly determined by realization capabilities of time unit (second) following the definition of the meter and which is in the uncertainty of second is in the order of  $1.5 \times 10^{-15}$  presently and, secondly, practical realization of temperature unit (Kelvin) as major disturbance parameter which influences the accuracy of dimensional measurements. However the actual precision of length measurements is limited by the other standard closely related with second that is the frequency standard - CH<sub>4</sub>-stabilized He-Ne laser. Satisfying the requirements of International Committee for Weights and Measures and the most common realization of primary length standard is He-Ne laser with the wavelength of  $\lambda = 632.99$  nm and relative standard uncertainty of  $2.5 \times 10^{-11}$  (Webb & Jones, 2004).

The bottleneck for calibration of material artifacts is determined by the capabilities of measurement systems to detect the boundaries/edges of geometrical features, requirements on environmental parameters as well as optical wave interferometry. In practice the lower limit for optical interferometry is in the range of  $10^{-7}$  and mainly restrained by variations of air refractive index which in turn can be computed by Edlen's formula contributing to the length measurement error in amount of  $5 \times 10^{-8}$  and might be considered as ultimate limit for measurements in air. On the other hand, vacuum interferometry might be an option where currently achievable expanded measurement uncertainties are below 5 nm measuring high quality artifacts in lengths below 1 m (Köning et al., 2007).

Typically for calibration of precision line scales and traceability to length standard optical comparators are used where optical or X-ray interferometry is used and such systems enable calibration of length standards with uncertainties ranging from several nanometers up to some tenth of nanometers when measured length is significantly longer in comparison with the wavelength of radiation source. However when geometrical dimensions of the measured structures are of the order of the wavelength the use of such technologies becomes quite complicated and therefore comparative methods using X-ray diffraction or scanning microscopes are more common, particularly for investigation of micro electro-mechanical system (MEMS) structures, lithography and similarly (PTB, 2003; Swiss Federal Office of Metrology, 2004).

While comparing these different methods it is necessary to consider not only the measurement range and accuracy requirements but also effectiveness of such systems relating to metrological network and needs of particular country, design costs, etc.

Although different calibration methods and equipment, see in Table 2, for dimensional measurement of a material artifact are used, currently, optical comparators combining properties of laser interferometers and optical microscopes represent the lowest relative uncertainty ( $U/L$ ) of dimensional measurements provided in a length calibration of a 1 m line scale, the relative expanded uncertainty (coverage factor  $k = 2$ ) is  $7 \times 10^{-8}$  m at 1 m.

Length calibration method	Measurement range, $\mu\text{m}$		Uncertainty $U=[a^2+b^2]^{1/2}$	
	Lower limit	Upper limit	a, nm	b, nm
Calibrated atomic microscope (NIST)	0.1	10	3	$2 \cdot L, L=0.003 \div 0.02 \mu\text{m}$
Laser diffractometer (NPL)	2	$1 \cdot 10^6$	3	$0.1 \cdot L, L, \text{mm}$
Scanning electronic microscope and laser interferometer (NPL)	0.29	50	0.025	$0.19 \cdot L, L=0.06 \div 61 \mu\text{m}$
Optical microscope laser interferometer and video microscope with CCD (MIKES)	0.05	10	10	$0.001 \cdot L, L, \mu\text{m}$
Combined X-ray and laser interferometer (NPL, PTB)	0.001	$1 \cdot 10^3$		$0.03, L < 10 \mu\text{m}$ $0.04, L < 100 \mu\text{m}$ $0.17, L < 1 \text{mm}$

Table 2. Comparison of different length calibration techniques (PTB, 2003; Swiss Federal Office of Metrology, 2004)

### 3. Length calibration systems

Length calibration problems are inevitably connected with mastering and research of modern interferometric displacement measurement technologies, error compensation technique as well as digital microscopy achievements. These problems are often determined by the embedded metrology needs that can be met only by developing novel systems that absorb recent scientific and technical findings and optimally comply with specific calibration requirements as well as by improving existing calibration systems open to complying with fundamental principles of precision engineering. Satisfying more demanding requirements ultimately drives the demand to trace/validate the products on their manufacturing line as well as performing the graduation line detection and position measurements in a short time.

The ability to provide reference measurements at such levels of uncertainty requires developments beyond the current state of the art in each of three areas:

- the physical artifact to be calibrated;



- the measuring machine to do the calibration
- the theoretical model of the systematic errors in measurement results arising from the interaction of the artifact and the measuring machine in the calibration process.

In addition, the three developments need be tied together in a measurement procedure that includes innovative measurement algorithms and methods.

For high precision line scale calibration and traceability to primary length standard optical comparators with movable optical line detection system or movable line scale are used. The interferometer is the instrument that transfers international standard of length into physical measurements. Digital measurement microscopes enable to perform precise positioning of length calibration systems, to estimate quality of line edges and precise location of lines. The system is placed in an air-conditioned environment.

The interferential comparator presented in Figs. 3, 4 was developed to calibrate both line graduation scales up to 3.5 m long and incremental linear encoders.

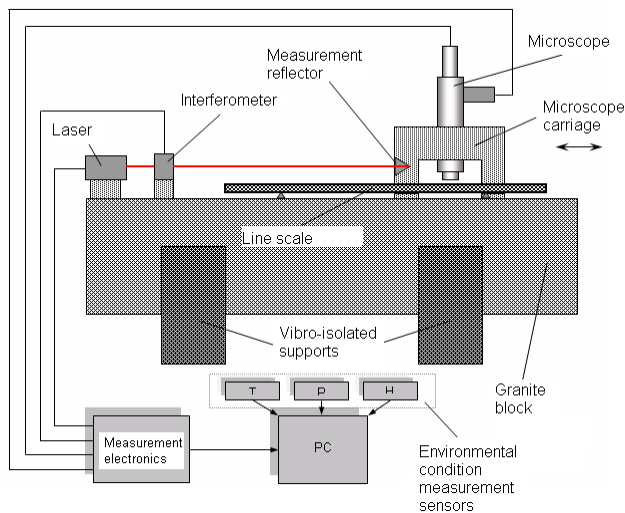


Fig. 3. Mechanical layout of the precision line scale comparator

The comparator consists of four main parts, namely the body of the machine, a laser interferometer, a translating system and a detecting apparatus. The body of the machine, which is made of granite surface plate, is used as the base of the machine and as a guide for the moving carriage. Measurement of the displacement of the carriage is realized by laser interferometer that consists of Zygo ZMI 2000 laser head and interferometer with the single-pass arrangement. The interferometer provides a resolution of 0.62 nm.

The comparator was designed to achieve expanded measurement uncertainties ( $k = 2$ ) down to  $7 \times 10^{-7}$  m ( $L = 1$  m) in dynamic regime. It enabled to trace the calibration of line scale of up to  $L \leq 3.5$  m long to the wavelength standard. The magnification and numerical aperture of the NIKON objective lens used was  $20\times$  and  $50\times$ , and numerical aperture - 0.4 and 0.55 respectively.

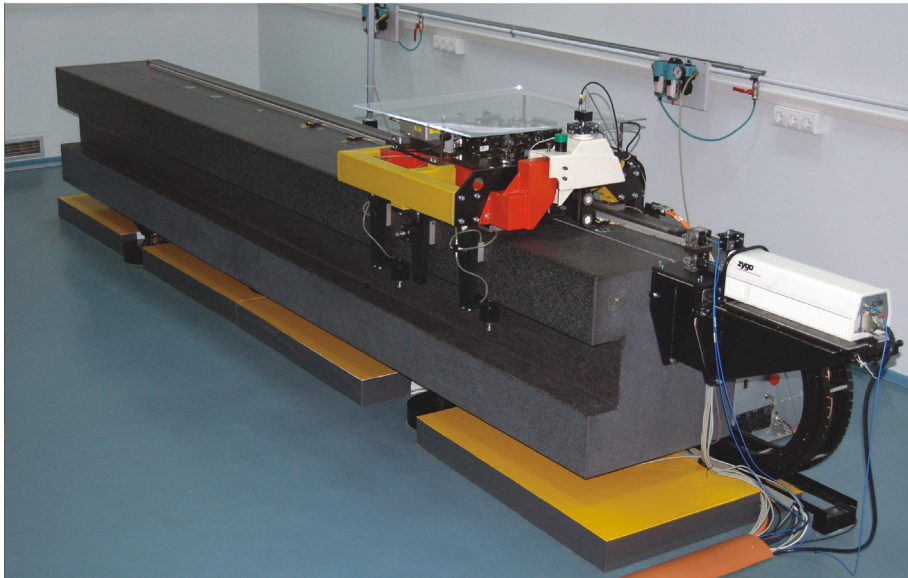


Fig. 4. Precision interferometer-controlled line scale comparator with charge-coupled device (CCD) microscope

A moving CCD microscope serves as structure localisation sensor for the measurements of line scales. The microscope on the carriage guided on aerostatic bearings is moved with a controlled velocity of 1–10 mm/s.

The graduation line distances are measured during continuous motion. Average profiles of the graduation lines are formed by summing picture element intensities of each row of the CCD. Line centre is calculated as weighted mean from intensity profile of a line.

Air pressure, temperature, humidity are on line accessed to determine the refraction index of the air by the Edlen's formula. The angular control loop - together with the numerical procedure - was applied to compensate and reduce the Abbe uncertainty contribution.

The whole calibration process and all operations of the system are controlled by the PC, which runs according to specific operation program that includes also the error compensation.

The measured performances confirm that investigated measurement system can operate reliable at velocities up to 6 mm/s without appreciable loss in measurement accuracy.

In order to examine the calibration process in real time, the experiments of the line scale calibration with a moving CCD microscope have been carried out in specific operating modes, and the accuracy of dynamic calibration was analysed.

### 3.1 Artifacts

Graduated length scales come in many forms, and are made in lengths from a few micrometers to over a several meters. Those longer than a meter or two are usually classified

as measuring tapes or rods. Many materials are used including steel, Invar, glass, glass-ceramics, silicon, and fused silica. Cross sectional shape can be rectangular, "H," modified "U" (flat bottom), or a modified "X" (Tresca). At present, the line scale interferometer is limited to graduation widths ranging from submicrometer to 100 mm, and spacings ranging from less than 1 mm up to 1025 mm. Spacings are generally measured from center to center of the graduations, but can also be measured from edge to edge.

Some devices that are not strictly linear scales are measured in the line scale interferometer. These include end standards in a size range (250 mm to 1000 mm) that can present measurement problems with laser interferometry.

Two dimensional patterns are measured by treating each row and column of graduations as an independent scale and, when possible, an estimate of orthogonality can be made by measuring the diagonals.

### 3.2 Error analysis of length calibration systems

Making measurements with nanoscale precision poses several major difficulties. Environmental fluctuations such as vibration or temperature change have a large effect at the nanoscale. For example, any external change to the large machines used in manufacturing microelectronics components will affect the creation of nanoscale features and their crucially important alignment to each other. The ability to measure these influences, and thereafter to minimize them, is therefore vital.

The error budget is an accuracy model of the machine in its environment expressed in terms of cause-and-effect relationships. It may involve random quantities such as seismic vibrations or deterministic quantities such as deflections due to gravity or weights of moving axes and payloads. The error budget helps identifying where to focus resources to improve the accuracy of an existing machine or one under development. It provides useful information for the specification of subsystem precision requirements to achieve an overall balance at the levels of difficulty.

The technical basis for the error budget rests on two assumptions:

- the total error in a given direction is the sum of all individual error components in that direction, and
- individual error components have physical causes that can be identified and quantified.

A practical difficulty arises because we generally cannot quantify the errors in complete detail especially at the design stage. Although an error may vary spatially and temporally, usually the only estimate will be a bounding envelope and perhaps an approximate frequency of variation.

In general, machine tools errors can be divided into two categories: systematic and random errors. Systematic errors can be described by deterministic mathematical and engineering models. Random errors are difficult to model and to compensate. In the calibration of line standards, several essential sources need to be identified and eliminated or corrected. The following error map, see Fig. 5, presented as cause and effect diagram, gives the sources of uncertainty classified by type and their origin (Jakštas, 2006).

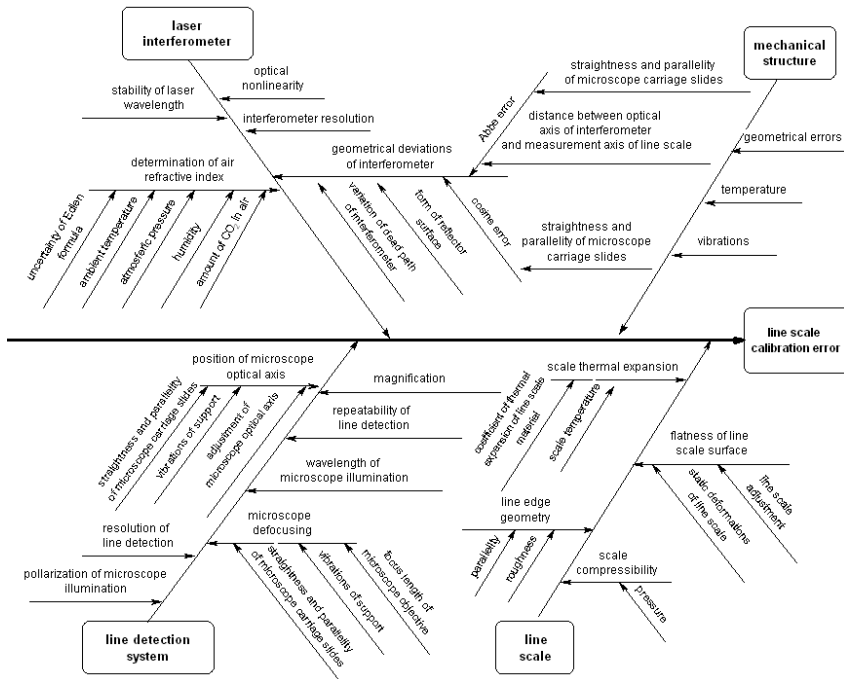


Fig. 5. Example of error model for line scale calibration system

Among the key factors that affect the accuracy of length calibration systems are geometric deviations and thermal effects on the comparator components and the scale. Mechanical limitations of calibration systems are featured by the whole complex of a mechanical system including error compensation circuits.

Thermally induced errors in measurement are an accuracy-limiting factor in the length metrology. Temperature, barometric pressure and humidity influence the refractive index of air and, thus, the wavelength of light; temperature also affects the length of the scale being measured. The philosophy of precise dimensional measurements is, therefore, to approximate thermal equilibrium near the reference temperature 20 °C for measurement condition. In order to calibrate especially long line scales some prerequisites have to be fulfilled as far as temperature influence is concerned:

- precise temperature measurements of scale and environment
- temperature stability over measurement time
- correction of small temperature drifts
- negligible thermal gradients
- temperatures close to reference temperature

Temperature problems are a widely assessed error source in precision measuring machines. While analyzing high precision calibration systems it is essential to evaluate an average volume temperature of some parts of a mechanical comparator as well as the temperature of the scale. Under real calibration conditions, temperature measurement is possible only at

certain points. The response time of temperature sensors is also rather long (from several seconds to several minutes). Therefore, fast temperature changes cannot be detected, and, consequently, the measurement uncertainty increases. The thermally induced errors in the long-stroke measuring machine are more significant because of their size and complexity. Due to high requirements for geometrical stability of calibration system, the temperature deformations caused by changes of several hundredth parts of Kelvin must be considered. The cause-and-effect relationships can be calculated in considerable detail using modern finite element (FE) analysis and empirical heat transfer formulas, but doing so requires considerable knowledge about the design and environment, see chapter 3.3.

The results of temperature measurements performed in the comparator laboratory (shown in Fig. 6) have revealed that maximum temperature deviations from the mean amount to about  $\pm 0.25$  K (during the whole period of system operation) and the main reasons that induce temperature gradients are laser source, computation and control unit, and thermally non-isolated part of the ground floor. The measurement of air temperature during the calibration procedure due to the introduced dynamic mode indicated much smaller temperature variances, and calculated standard deviation was 0.034 K. For precision calibration of the line scales not only the closeness of laboratory temperature to nominal 20 °C temperature is of crucial importance, but also temperature stability (deviations) during calibration procedure. Therefore, a period of temperature stabilization that can last for 10-14 hours is necessary in order to accomplish high precision calibration procedures.

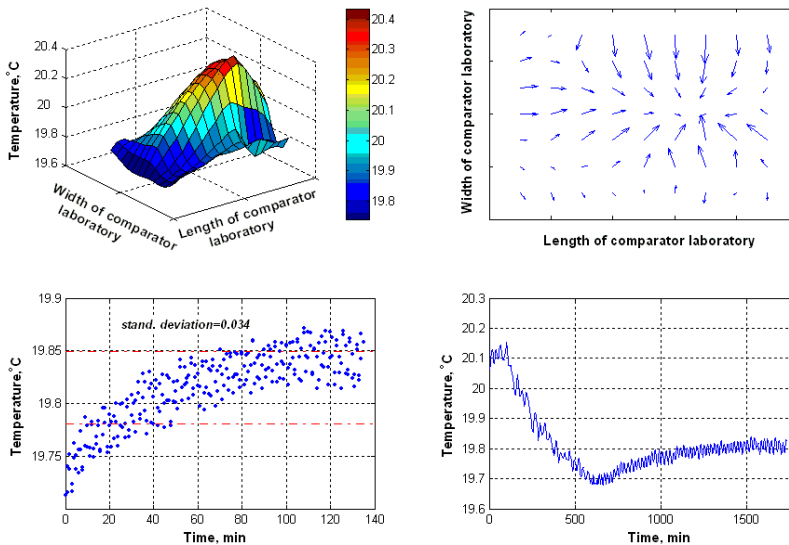


Fig. 6. Temperature measurements in the laboratory of length calibration system

### 3.3 Modeling

Error-related problems specific to length measurements are caused primarily by geometrical and thermal deviations of the comparator components and the scale. One of critical tasks in

the dynamic calibration of the scales is microscope image acquisition, i.e. bringing the specimen into focus before taking an image and measuring any feature. Due to imperfections of the stage, i.e. inaccuracy of the motion of the scanning mechanism and vibrations, the microscope slide is not perfectly perpendicular in regard to the optical axis of the imaging system. Measuring even a slightly vibrating structure with any degree of accuracy is prone to an error with an optical microscope. Any deviation in the distance of microscope lens with respect to the scale - for example, when the surface is in motion during the data acquisition process - introduces measurement errors. The magnitude of the resulting error can range from a few nm to several  $\mu\text{m}$  depending on the magnitude of such disturbances and the measurement setup. The most common error associated with small vibrations is the error in detection of the line scale graduation.

Structures of precision length measuring machines are often too sophisticated to be modelled precisely by applying simple methods. Therefore complex models as well as their analysis tools are needed in order to perform qualitative and quantitative description and analysis of determinants of the precision length calibration process.

In this work finite element method (FEM) techniques were used for behavior simulation of the comparator carriage, CCD microscope and calibrated line standard itself under the influence of dynamic and thermal factors, like variations of environmental temperature, vibrations in the structure caused by seismic excitation.

The state-of-the-art FE technique was applied in order to evaluate the possible influence of dynamic and thermal factors upon the inaccuracies of measurement. Two basic physical phenomena were of interest:

1. displacements of the structure as a consequence of applied dynamic excitations, as well as, non-homogeneities of the temperature field;
2. heat transfer inside of the structure caused by an external temperature field.

All necessary aspects of the dynamic behavior of the comparator can be investigated by employing small displacement elastic structural models as (Jakštas et al., 2006)

$$[M]\{\ddot{U}\} + [C]\{\dot{U}\} + [K]\{U\} = \{F(t) + Q(t)\} \quad (2)$$

$[K]$ ,  $[M]$ ,  $[C]$  are stiffness, mass and damping matrices of the structure;  $\{F\}$  is nodal vector of external excitation forces;  $Q(t)$  is nodal force vector caused by the temperature propagation effect;  $\{U\}$ ,  $\{\dot{U}\}$ ,  $\{\ddot{U}\}$ , are nodal displacement, velocity and acceleration vectors.

FE models enable us to simulate all 3D displacement or vibration patterns of the structure. Vertical vibrations of the microscope may lead to defocusing. Vibrations in the direction of motion may cause detection errors in determining positions of graduation lines.

In practice the excitation vectors  $F(t)$  caused by external dynamic effects or by moving parts of the structure are not known explicitly, but often are subjected to external excitation propagating through the base and supports of the comparator structure. A spectrum analysis is one in which results of a modal analysis are used with a known spectrum to calculate displacements and stresses in the structure modeled. The model is capable of predicting the system's behavior under thermal load and enables us to investigate the

thermo-mechanical processes in the system, by taking into account both static and dynamic disturbances and parameter deviations.

The dynamic response of the comparator as a result of seismic excitation was investigated by employing small displacement elastic structural models (Jakštas, 2006; Jakštas et al. 2006). Modeling of seismic excitations in the comparator structure has shown that maximum displacements are expected at the bottom plane of the microscope objective and can amount more than 100 nm.

Modal analysis of the spatial carriage and microscope deviations induced by seismic excitations as well as those caused by operation of the carriage drive vibrations has shown that dynamic factors may contribute significantly to the calibration uncertainty budget. To minimize, in particular, the vibrations of the measurement reflector the construction of the carriage structure was optimized and drive-originated were vibrations reduced.

Extensive investigations were accomplished to both reduce the dynamically induced deviations originated by the dynamic excitations of the mechanical structure and optimize the comparator design. The precision measurements were performed in order to evaluate the impact of small vibration on performance of the line scale calibration process. The experimental results revealed, in particular, that the sample standard deviation of the drive-induced relative displacements between the moving reflector of the interferometer and the measurement point of the microscope may reach  $0.662\ \mu\text{m}$  (at calibration speed  $3\ \text{mm/s}$ ); they were considerably reduced by optimization the carriage structure and elimination the undesirable modes of vibration.

Measurement results depicted in Fig. 7 show an improvement of systems stability using the optimized microscope carriage structure (blue line). The sample standard deviation of the drive-induced relative displacements of measurement mirror was reduced from  $0.178\ \mu\text{m}$  down to  $0.054\ \mu\text{m}$  (Kaušinis et al., 2009).

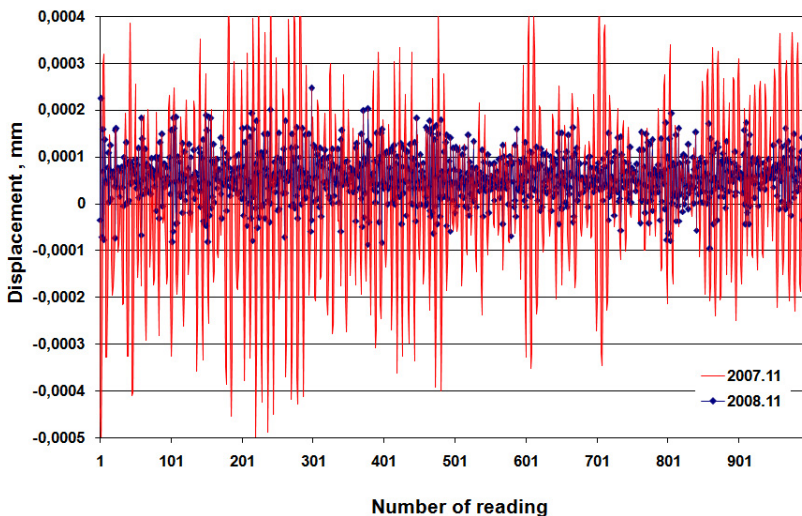


Fig. 7. Comparison of vibrations of the interferometer measurement mirror

Temperature problems are a widely assessed error source in precision measuring machines. While analyzing high precision calibration systems it is necessary to evaluate an average volume temperature of some parts of a mechanical comparator as well as the temperature of the scale. Under real calibration conditions, temperature measurement is possible only at certain points. The response time of temperature sensors is also rather long (from several seconds to several minutes). Therefore, fast temperature changes cannot be detected, and, consequently, the measurement uncertainty increases. The thermally induced errors in the long-stroke measuring machine are even more significant because of their size and complexity. Due to high requirements for geometrical stability of calibration system, the temperature deformations caused by changes of several hundredth parts of Kelvin must be considered.

The cause-and-effect relationships can be calculated in considerable detail using modern FE analysis and empirical heat transfer formulas, but doing so requires considerable knowledge about the design and environment.

The impact of temperature on the mechanical deformation of the line scale can be simulated in several ways:

1. temperature values at certain points of the construction can be detected experimentally and used for calculation;
2. temperature field can be calculated depending on the assigned non-homogeneity of the environment temperature by taking into account heat convection processes between the parts of the structure and its surrounding.

One of precarious temperature disturbances is the heat spread out by the CCD camera of the measuring microscope. As the steady-state temperature under the operating conditions is known, the thermal expansion process can be modeled by using the FE simulation, and the temperature values can be found at all points of the microscope structure. Having the temperature values obtained, the displacements due to thermal expansion can be calculated at any point of the structure.

The equation of the structure heat balance reads as follows:

$$[C]\{\dot{T}\} + [K_{Th}]\{T\} = \{S_{\infty}\} \quad (3)$$

where  $[C]$  - matrix of thermal capacity,  $[K_{Th}]$  - matrix of thermal conductivity,  $\{S_{\infty}\}$  - nodal vector of heat sources of the element determined by the heat exchange over the surface of the body.

The solution presents the nodal temperature values, which are further used as loads in the problem of thermal expansion of the structure as:

$$[K]\{U\} = \{Q\} \quad (4)$$

where  $[K]$  - stiffness matrix of the element;  $\{Q\}$  - vector of nodal forces determined by temperature loads.

The FE computational model of the structure was set up, in which the temperatures of the structure and the ambient air could be calculated. The model is based on the coupling of the following physical phenomena:



1. Heat transfer by the ambient air due to its thermal conductivity;
2. Convective heat transfer (due to the motion of the air);
3. Heat exchange between the air and comparator structure;
4. Heat transfer by comparator structure due to its thermal conductivity;
5. Formation of deformations in comparator structure due to the non-homogenous thermal field generated in it.

In the computational model phenomena 1-4 have been described by means of ANSYS (FLOTTRAN) element FLUID142. The element can be used under two different conditions:

- liquid (gas) dynamics described by the continuity equation, the advection-diffusion equation and the ideal gas state equation;
- thermal conductivity in the solid described by the thermal conductivity equation.

The model is capable to predict the system's behavior under thermal load and enables us to investigate thermo-mechanical processes in the system and facilitates finding proper structural solutions to reduce the impact of thermal load on the calibration accuracy.

Displacements in the structure caused by the calculated temperatures field are depicted in Fig. 8.

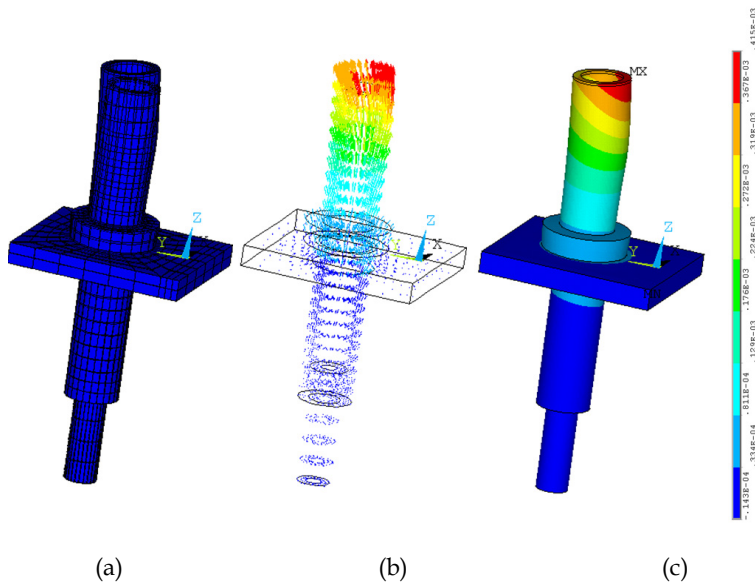


Fig. 8. Displacements in the structure caused by the calculated temperatures field; CCD camera is fixed to a side of the microscope: (a) deformation of the construction; (b) vectors of nodal displacements; (c) deformation of the structure

Vector  $[Q]$  in equation (1) can be easily determined if the temperature field inside of the structure is known. If only the surrounding transient temperature field is known, the temperatures inside of the structure can be obtained by solving the thermal conductivity problem

$$[C_{th}]\{\dot{T}\} + [K_{th}]\{T\} = \{S_{\infty}\} + \{Q_{th}\} \quad (5)$$

where  $[C_{th}]$ ,  $[K_{th}]$  - thermal capacity and conductivity matrices of the structure,  $\{S_{\infty}\}$  - nodal vector of heat sources determined by heat convection across the surface of the structure.

In order to determine distribution of the temperature fields around the CCD camera, temperature sensors were arranged, and temperature distribution measured at CCD cut-off, warming up, and steady operating conditions.

Calibration error caused by a thermal CCD camera impact under steady-state calibration conditions is of a random character and in real-time it cannot be compensated by mathematical methods. Within the experiments conducted an estimate of the variance of this error amounts to  $\pm 0.23 \mu\text{m}$  at the 95% probability level (Barauskas et al., 2011).

### 3.4 Line scale calibrations

High-precision measurements of line scales basically apply two main modes of calibration: static and dynamic. Currently, static line detection systems are predominantly used in metrology institutes and calibration laboratories worldwide. The static method is potentially more accurate but somewhat slow, whereas the dynamic method offers taking advantages of scale calibration in terms of speed, accuracy and throughput. It also allows the construction simplification, because high-precision settling of the moving scale or microscope is unnecessary, and the measurement process is less influenced by ambient environmental conditions. On the other hand, the dynamic method encounters difficulties induced by measurement speed fluctuations, time delays, noise and vibrations especially during the graduation line detection. Mechanical limitations for the dynamic mode are featured by the whole complex of mechanical system including error compensation circuits.

As the dynamic calibration process is to be examined in real time, the experiments of the line scale calibration with both a slit and moving CCD microscope have been carried out in specific operating modes, and the dependence of the accuracy of dynamic calibration vs. speed has been studied.

The calibration experiments were performed that intended to document current capabilities to carry out line scale calibrations on high quality graduated scales made of low thermal expansion substrates. The line scale standard made of the glass ceramic Zerodur was available for calibration purpose from PTB. The dimensions of the scale are 230 mm in length, 25 mm in width and 14 mm in height. The graduation represents a total length of 200 mm and consists of line structures with 1 mm length and  $2.5 \mu\text{m}$  width. The line structures are reflecting on transparent substrates. The measurand that was determined on the line scale standard is the deviations from the nominal lengths for 1 mm lines (1 mm pitch). Fig. 9 shows the deviations from the nominal positions for the weighted mean, calculated on the basis of the set of 6 independent measurement runs taken at the microscope carriage speed 3mm/s. The environmental chamber and scale temperatures were held within  $\pm 0.05 \text{ }^{\circ}\text{C}$  during the measurements. The positions of the line are corrected for the influence of the temperature deviation from  $20 \text{ }^{\circ}\text{C}$  and pressure deviations from 1013.25 hPa.

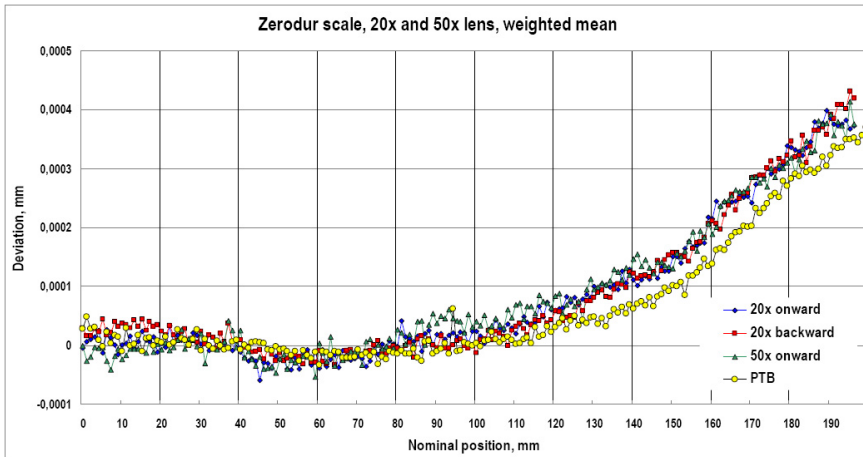


Fig. 9. Calibration results on Zerodur scale, 200 mm graduation, 1 mm step.

It has been demonstrated that the capabilities of a newly developed comparator are close to the calibration capabilities of analogous long scale calibration systems in the other countries, available from BIPM key comparison database, see Fig. 10, and still can be improved first of all by embedding an automatic line focusing system and tightening tolerances of ambient conditions in the laboratory.

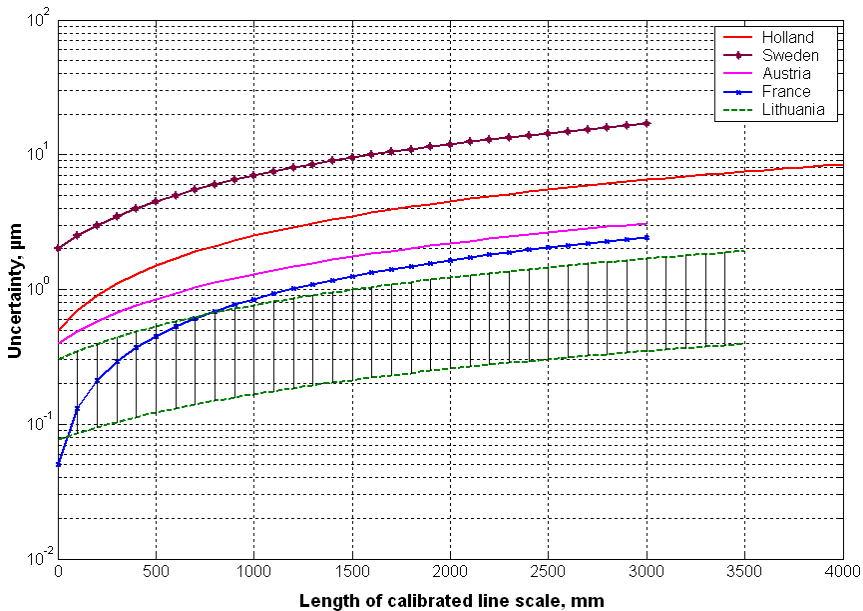


Fig. 10. Comparison of long line scale calibration capabilities

## 4. Conclusions

Precision measurements based on the length standard are very important, and their significance is growing with the advancement of science and industrial technology. Particularly important is product quality improvement in production processes. The development of metrology standards follows the needs of technology.

Diversity of tools and measurement techniques in dimensional metrology require significantly more mature set of standards for characterization of the measurement process used to obtain measurement data, e.g. what sampling strategy was used, what filtering was applied and what measurands definition was applied. It will also expect more know-how on the user level of the metrology involved in operating modern measurement equipment.

In industrial metrology, several issues beyond accuracy constrain the usability of metrology methods. These include, among other factors, the speed with which measurements can be accomplished on parts or surfaces in the process of manufacturing, and the ability of the measurement system to operate reliably in a manufacturing plant environment considering temperature, vibration, dust, and a host of other potential hostile factors.

The relevance and necessity of addressing the problem of precision and high-speed line scale calibration is primarily driven by the rapid increase of demands on calibration efficiency of precision scales. Considerably higher precision and efficiency requirements are set for new systems, besides, it is aimed at traceability of precision line scale parameters during manufacturing process in the technological line, and calibration process should be as short as possible.

This chapter comprehends a synopsis and analysis of literature and existing scientific and technical solutions of precision 1D length calibration. It addresses also the problems of the research and development of an interferometer-controlled comparator that is operated in non-ideal environmental conditions and enables to trace the calibration of line scale of up to  $L \leq 3.5$  m long to the wavelength standard.

The analysis and research results represent both systematic methodology and knowledge base for evaluation the length calibration accuracy that involve current and new technologies, and can be applied gradually in various precision machinery and instrumentation.

## 5. References

- Barauskas, R.; Kasparaitis A.; Kaušinis S. & Lazdinas R. (2011). Temperature fields, exchanges and deformations of a precise comparator structure, *Mechanika*, No.17, Kaunas, Lithuania, pp. 279-283, ISSN 1392-1207
- Beers, J.S. & Penzes W.B. (1999). The NIST Length Scale Interferometer, *Journal of Research of NIST*, Vol.104, pp.225-252
- Bobroff, N.H. (1993). Recent advances in displacement measuring interferometry, *Measurement Science Technology*, Vol.4, pp. 907-926
- Bosse, H. & Flügge, J. (2001). Requirements and recent developments in high precision length metrology, *Proceedings of the 159. PTB-Seminar*, 2001, 180 p.
- Bureau International des Poids et Mesures. (April 2003). Evolving needs for Metrology in Trade, Industry and Society and the Role of the BIPM, 142p.
- Chapman, M. (2002). Heterodyne and homodyne interferometry, Renishaw

- DTI/NMDS (January 2002). NMS Science & Technology Programme 1999-2002: Length, UK
- DTI/NMDS (October 2002). NMS Science & Technology Programme in Length Metrology (October 2002 – September 2005), UK
- Farago, F.T. & Curtis M.A. (1994). Handbook of Dimensional Measurements, 3rd Edition, New York: Industrial Press Inc., 580p.
- Israel, W.; Tiemann, I.; Metz, G.; Yamaryo, Y.; Maeda, F. & Shimomura, T. (2003). An international length comparison at an industrial level using a photoelectric incremental encoder as transfer standard, *Precision Engineering*, Vol. 27, pp. 151-156
- Jakštas, A.; Kaušinis, S.; Barauskas, R. (2006). Modeling of precision line scale comparator, *Mechanika*, No. 2(58), Kaunas: Technologija, Lithuania, pp. 57-62, ISSN 1392-1207
- Jakštas A. (2006). Investigation of precision length calibration systems, Summary of Doctoral Dissertation, Kaunas University of Technology, Kaunas, Lithuania
- Jakštas, A.; Kaušinis, S.; Barauskas, R.; Kasparaitis, A. & Barakauskas, A. (2011). Investigation of dynamics-induced errors of long line scale calibration systems, *Measurement* Vol.44, Issue 5, pp. 976-987, ISSN 0263-2241
- Kaušinis, S.; Jakštas, A.; Flügge, J.; Barauskas, R. (2004). 1D high precision interferometer-controlled comparator, *Proceedings of the 9th Biennial Baltic Electronics Conference*, October 3-6, 2004, Tallinn University of Technology, Estonia, 2004, pp. 195-198, ISBN 9985594622
- Kaušinis, S.; Jakštas, A.; Barauskas, R.; Kasparaitis, A. (2006). Investigation of dynamic properties of line scale calibration systems, *IMEKO XVIII World Congress and IV Brazilian Congress of Metrology*, September 17-22, 2006, Rio de Janeiro, Brazil
- Kaušinis, S.; Kasparaitis, A.; Barakauskas, A.; Barauskas, R.; Jakštas, A.; Kilikevičius, A. (2009). Line scale calibration in non-ideal measurement situation, *Solid State Phenomena: Mechatronics Systems and Materials III*, selected, peer reviewed papers from the 4th International Conference: Mechatronics Systems and Materials (MSM 2008), Bialystok, Poland, 14-17, July, 2008. Uetikon-Zuerich: Trans Tech Publications Ltd. I, Vol. 147-149, pp. 682-685, ISSN 1012-0394. 2009
- Kaušinis, S.; Barakauskas, A.; Barauskas, R.; Jakštas, A.; Kasparaitis, A. (2009) Reducing dynamically-induced deviations for line scale calibration in non-ideal measurement situation, *XIX IMEKO World Congress: Fundamental and Applied Metrology*, September 6-11, 2009, Lisbon, Portugal, pp. 1971-1974, ISBN 9789638841001
- Köning, R.; Flügge, J. & Bosse H. (2007). Achievement of sub nanometer reproducibility in line scale measurements with the nanometer comparator, *Proc. SPIE 6518*, 65183F.
- PTB (2003). WDM-7 Preliminary comparison on nanometrology according to the rules of CCL key comparison Nano3, Line Scale Standards, Final Report, Braunschweig, 119p.
- Quinn, T.J. (2003). Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001), *Metrologia*, Vol.30, pp.103-133
- Swiss Federal Office of Metrology (2004). WDM-7: Preliminary comparison on nanometrology according to the rules of CCL key comparisons Nano4, 1D gratings, Final report, Wabern, 30. November 2004, 34p.
- Swyt, D.A. (2001). Length and Dimensional Measurements at NIST, *Journal of Research of NIST*, Vol.106, pp. 1-23
- Webb. C.E. & Jones J.D.C. (2004). Handbook of laser technology and applications Volume 2 - Laser Design and Systems, Institute of Physics, Bristol, ISBN 0750309636, 625p.
- Yim, N.B.; Eom C.I.; Kim S.W. (2000). Dual mode phase measurement for optical heterodyne interferometry, *Measurement Science Technology*, Vol.11. pp.1131-1137

# Dimensional Metrology and Frequency Sweeping Interferometry

Alexandre Cabral, José Manuel Rebordão and Manuel Abreu

*Faculty of Sciences, Physics Department and  
Centre for Astronomy and Astrophysics of the University of Lisbon, University of Lisbon,  
Portugal*

## 1. Introduction

Absolute distance metrology is needed for a wide gamut of applications with different ranges and resolutions. One good example is space missions requiring independent satellites working cooperatively: the management of such formations requires several levels of metrology to keep the formation coherent and to enable guidance and navigation of the complete formation (Calvel et al., 2004). Beyond a certain level of accuracy, at the end of the metrology chain, optical metrology becomes mandatory to achieve the required high accuracy.

Optical metrology is normally used to describe the measurement of some physical parameter using optical methods. Almost all of the technical fields we can think of make use of optical metrology. A good example is the use of interferometry to achieve high accuracy in length measurements. Optical interferometry relies on the phenomenon of interference between light waves to make extremely accurate measurements. In 1887, Michelson, along with physicist Edward W. Morley, set up an experiment to determine if the speed of light was dependent on the speed of the observer. According to the accepted theory of that time, light had to propagate in a medium called the ether. The motion of Earth travelling through the ether would affect the fringe pattern on Michelson's interferometer, because it would take the light longer to travel over one arm than the other arm. When the interferometer was rotated through  $90^\circ$ , the fringes would shift if the speed of light was not constant. In this most celebrated null experiment in the history of science, Michelson and Morley observed no changes in the fringes after many repetitions. The speed of light appeared to be constant, regardless of the speed of its source, a fact that was later explained by Einstein's theory of relativity.

The "failure" of the Michelson-Morley experiment was the seed of the new science of interferometry. The improved device used in the Michelson-Morley experiment was one of the first interferometers, and Michelson, who had always an intense passion for accuracy in measurement, quickly recognized his invention's potential for highly accurate instrumentation.

Since its debut, the interferometer (Fig. 1) has undergone a number of modifications and has been specialized for use in a variety of fields. In 1895, Michelson used it to measure the International Prototype Metre in Paris in units of wavelength (Michelson & Benoit, 1985).

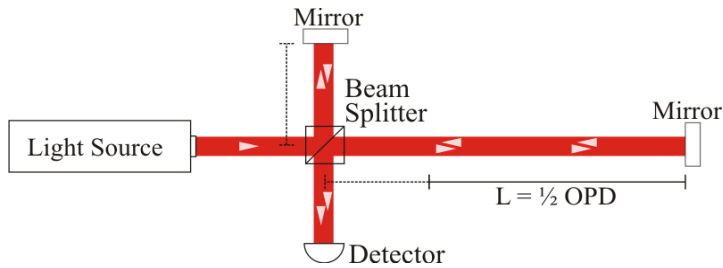


Fig. 1. Setup of a Michelson interferometer.

Michelson's contributions to interferometry, from 1880 to 1930, dominated the field to such extent that optical Interferometry was regarded for many years as a closed chapter. However, the last few decades have seen an explosive growth of interest in interferometry due to several new developments, mainly associated to the light source. The improvement of frequency stabilised lasers, of frequency tunable lasers and, lately, the appearance of femtosecond lasers have encouraged the use of the interferometer in dimensional metrology. In addition, nanometre-level interferometry with low-coherence sources has also created a plethora of additional metrological applications and technological developments. Interferometry is still alive in the XXI century.

Coherent absolute distance metrology is one of the most interesting techniques for dimensional metrology. Without movement of the mirror that define the Optical Path Difference (OPD) in the interferometer, measurements are made without ambiguity, by using either one or several synthetic wavelengths resulting from the beating of two or more wavelengths (multiple wavelength interferometry) or, in the case of frequency sweeping interferometry, from an optical frequency sweep.

Frequency Sweeping Interferometry (FSI) based sensors are relatively simple devices and can fulfil an important role in dimensional metrology. In addition, their parameterisation flexibility allows various tradeoffs to be performed, either technology driven or application related.

The generation of the synthetic wavelength in FSI is based on optical frequency sweeping the laser source within a given sweep range. As frequency sweeps, detection electronics counts synthetic wavelength maxima (temporal "synthetic fringes") without ambiguity. Its sensitivity to variations of distance (drift) during the sweep limits the maximum resolution to a few micrometres. FSI does not require stabilised or well known laser sources and relies only on a tunable laser and a frequency sweep range measurement subsystem (for example, based on a Fabry-Pérot interferometer). FSI sensor complexity can be tuned to system specifications – when requirements are modest, FSI complexity is reduced accordingly.

In the last years, FSI has seen a considerable development and several solutions have been presented. The most important are based only upon a tunable laser and a frequency measurement sub-system, systems built around both a reference and a measurement interferometer, or based on the use of an additional stabilized laser, some developed for high accuracy under special controlled laboratory conditions, others driven by robustness in

uncontrolled environments (Kinder & Salewski, 2002; Cabral & Rebordão, 2005; Swinkels et al., 2005; Yang et al., 2005). An example is the Dual FSI concept, conceived to overcome the decrease in accuracy as the measurement range increases, as a consequence of the uncertainty propagation in the synthetic wavelength measurement. To overcome this problem, in order to maintain the low complexity associated to short distances, in Dual FSI the measurement process for larger ranges is reduced to the close range case, by increasing the reference arm with a long reference fibre and using an ancillary interferometer to measure (calibrate) continuously the fibre length.

## 2. Absolute Distance Measurement by optical interferometry

Absolute Distance Measurements (ADM) can be performed by several optical methods. Generally, one has non-interferometric methods for large ranges with moderate accuracies and interferometric methods for high accuracies and moderate ranges, limited by the coherence length of the source. ADM can also be performed on the basis of low-coherence or white light interferometry but, in this case, the measuring range is limited to a few millimetres. Lately, the measuring range of interferometric methods increased due to the increase of the coherence length and the development on new coherent sources.

ADM interferometric techniques are typically based on the use of multiple or time varying wavelength sources. Single wavelength interferometry is a powerful tool in displacement measurement when high accuracy is required, however, as a major drawback, is only applicable to relative measurements. To measure an absolute distance it is necessary to use incremental interferometry: to measure a displacement it is necessary to fix a reference position and to carefully displace a suitable reflector from that position to the final one, and basically counting each interferometer fringe event. This must be done avoiding any optical misalignment that would result in a loss of the interference signal and therefore of the measure.

Multiple Wavelength Interferometry (MWI) is a technique where the use of (at least) two different wavelengths allows the generation of a synthetic wavelength much longer than the two individual optical wavelengths, actually increasing the non-ambiguity range (NAR) for interferometry (Dändliker et al., 1988; Dändliker et al., 1995). In MWI, each synthetic wavelength (one or several) is generated using two different and very accurately known optical wavelengths. By selecting a small synthetic wavelength it is possible to achieve high resolution (at the micrometre level or even smaller). However the intrinsic ambiguity of the measurement limits the measurement range. To overcome the ambiguity, a chain of increasing synthetic wavelengths is generated (or a combination with another sensor) to cover the required range (Salvadé et al., 2000). For large ranges, several wavelengths are required and MWI can become indeed a very complex solution.

In a first approach, FSI is equivalent to MWI. In FSI, the generation of the synthetic wavelength is based on frequency sweeping the laser source within a given sweep range. The technique is not new, has strong similarities to radar, dating back to the 80's (Kikuta et al., 1986), but it was not studied extensively until the development of tunable lasers and the emergence of External Cavity Diode Lasers (ECDL) (Hecht, 2001). As frequency sweeps, detection electronics counts synthetic wavelength fringes (temporal "synthetic fringes")



without ambiguity, thus making it particularly interesting for large measurement ranges (Thiel et al., 1995; Stone, 1999; Edwards et al., 2000; Coe et al., 2004; Cabral & Rebordão, 2005; Swinkels et al., 2005). While in Double Wave interferometry (DWI) – which is a particular case of MWI – we have a fixed synthetic wavelength, while in FSI, as the frequency is being swept, the value of the synthetic wavelength is decreasing down to a value defined by the total sweep range.

In contrast to MWI, FSI does not require independent stabilized and well known laser sources and relies only on a tunable laser and a frequency sweep range measurement subsystem, normally based on a Fabry-Pérot interferometer (FP).

As in DWI, the synthetic wavelength  $\Lambda$  is inversely proportional to the frequency sweep range  $\Delta\nu$  (see next section). While in MWI we measure only the fractional part of the synthetic wavelength fringe, in FSI both the fractional and integer number of synthetic fringes can be measured. Thus, the absolute value of the OPD (between the two arms of a Michelson interferometer) will be determined without ambiguity.

It is not possible to elaborate an objective comparison between different techniques without an a priori definition of the sensor requirements: maximum range, accuracy, technological complexity, reliability, cost, etc. Nevertheless, a trade-off between the two basic methods, FSI and MWI, can be done in a qualitative way, considering the three main parameters: measurement distance, measurement uncertainty and overall complexity of the sensor subsystem.

FSI main advantage is the capability to perform large range measurements due to the non-ambiguity nature of the technique, limited only by the coherence length of the tunable laser (a few hundred metres). FSI main drawback is the sensitivity to drift that, even with a compensation method, will limit the accuracy at the micrometre level. With an affordable complexity in the frequency sweeping (i.e. synthetic wavelength) measurement sub-system, for large distances, FSI can achieve a relative accuracy around  $10^{-5}$ . With a high degree of complexity (locking the laser to a resonant cavity) accuracies can be improved by one or two orders of magnitude.

### 3. Frequency Sweeping Interferometry

In the previous section it was mentioned that the intensity detected in a Michelson interferometer changes with the OPD change for a fixed laser frequency. If the frequency changes in time, the detected intensity will change not only due to the OPD change but also due to the frequency change.

FSI is based in the variation of the optical frequency  $\nu$ , or angular frequency  $\omega = 2\pi\nu$ , of a tunable laser in a Michelson interferometer. In this situation, the frequency no longer represents the periodical property of the wave, and in this case, the phase component is a non-linear function of time. Defining the instantaneous angular frequency as the derivative of the phase:

$$\omega(t) = \frac{d\phi(t)}{dt} \quad (1)$$

the phase is given by:

$$\phi(t) = \int_0^t \omega(t) dt + \phi_0 \quad (2)$$

where  $\phi_0$  is the initial phase of the light source. The electric-field amplitude (not considering its spatial variation) is then described by:

$$E(t) = A(t)e^{i\phi(t)} \quad (3)$$

If two optical waves are derived from the same coherent light source, the angular frequency changing linearly in time (during the measurement period) but travel along different paths and are recombined at a point in space to interfere, the detected signal will change in time, unlike the case of a fixed frequency. In a Michelson interferometer, the angular frequency of the reference wave within the time interval  $t \in [0, \Delta t]$  is given by:

$$\omega_R(t) = \Omega \cdot t + \omega_0 \quad (4)$$

where  $\omega_0$  is the angular frequency at the beginning of the frequency sweep, and  $\Omega$  is the angular frequency sweep rate. If the angular frequency variation has a linear behaviour, the value of the angular frequency sweep rate is given by:

$$\Omega = 2\pi \frac{\Delta\nu}{\Delta t} \quad (5)$$

where  $\Delta\nu$  is the frequency sweep range. The phase of the reference wave is given by:

$$\phi_R(t) = \frac{1}{2}\Omega \cdot t^2 + \omega_0 \cdot t + \phi_0 \quad (6)$$

The wave function of the reference electric field is described by:

$$E_R(t) = A_R \cdot e^{i\left[\frac{1}{2}\Omega t^2 + \omega_0 t + \phi_0\right]} \quad (7)$$

where  $A_R$  is the amplitude of the reference wave (considered to be constant in time). The signal wave, travelling in the other arm of the interferometer, can be described similarly:

$$\omega_S(t) = \Omega \cdot (t + \tau) + \omega_0 \quad (8)$$

$$\phi_R(t) = \frac{1}{2}\Omega \cdot (t + \tau)^2 + \omega_0 \cdot (t + \tau) + \phi_0 \quad (9)$$

$$E_S(t) = A_S \cdot e^{i\left[\frac{1}{2}\Omega(t+\tau)^2 + \omega_0(t+\tau) + \phi_0\right]} \quad (10)$$

where  $A_S$  is the amplitude of the signal wave (considered to be constant in time) and  $\tau$  is the delay time of the signal wave with respect to the reference wave. The delay is related to the OPD in the interferometer by:

$$\tau = \frac{n \cdot OPD}{c} \quad (11)$$

where  $n$  is the refractive index of the optical propagation medium. When these two waves interfere, the intensity of the resulting electrical field can be written as:

$$\begin{aligned} I &= \langle |\mathbf{E}_R|^2 \rangle + \langle |\mathbf{E}_S|^2 \rangle + \langle \mathbf{E}_R \cdot \mathbf{E}_S^* \rangle + \langle \mathbf{E}_R^* \cdot \mathbf{E}_S \rangle \\ I(\tau, t) &= |A_R|^2 + |A_S|^2 + 2(A_R \cdot A_S) \cos\left(\Omega \cdot \tau \cdot t + \omega_0 \cdot \tau + \frac{1}{2} \cdot \Omega \cdot \tau^2\right) \end{aligned} \quad (12)$$

or equivalently by:

$$\begin{aligned} I(\tau, t) &= I_R + I_S + 2\sqrt{I_R I_S} \cdot \cos\left(\Omega \cdot \tau \cdot t + \omega_0 \cdot \tau + \frac{1}{2} \cdot \Omega \cdot \tau^2\right) \\ &= I_0 \left[ 1 + V \cdot \cos\left(\Omega \cdot \tau \cdot t + \omega_0 \cdot \tau + \frac{1}{2} \cdot \Omega \cdot \tau^2\right) \right] \end{aligned} \quad (13)$$

where  $I_R$  and  $I_S$  are the intensities of the reference and signal wave,  $I_0$  is the average intensity and  $V$  is the visibility (or contrast) of the signal:

$$V = \frac{2\sqrt{I_R I_S}}{I_R + I_S} \quad (14)$$

The relative importance of the different terms of the cosine argument in Eq.(13) depends on  $\tau$ ,  $\Omega$  and  $\omega_0$ . The most relevant parameter is the delay  $\tau$ . This value is limited by the coherence length of the tunable laser. For a coherence length of a few hundred metres,  $\tau < 1 \mu\text{s}$ , the value of  $\frac{1}{2}\Omega \cdot \tau^2$  will be several orders of magnitude lower than the other two terms (a typical value for  $\Omega$  is  $10^4$  GHz/s). It is thus conceivable to neglect the term  $\frac{1}{2}\Omega \cdot \tau^2$  in Eq.(13), and simplify the equation:

$$I(\tau, t) = I_0 \left[ 1 + V \cdot \cos(\Omega \cdot \tau \cdot t + \omega_0 \cdot \tau) \right] \quad (15)$$

If we rewrite Eq.(15) considering the optical frequency and the OPD to describe the optical waves and the interference phenomena, we obtain (from Eq.(5) and (11)):

$$I(OPD(t), t) = I_0 \left[ 1 + V \cdot \cos\left(2\pi \left( \frac{\Delta\nu}{c} \cdot n \cdot OPD(t) \cdot \frac{t}{\Delta t} + \frac{n \cdot OPD(t)}{\lambda} \right) \right) \right] \quad (16)$$

where  $OPD(t)$  represents the OPD variation during the sweep and  $\lambda$  is the optical wavelength (in vacuum) at the beginning of the sweep. Defining a synthetic wavelength  $\Lambda$ , inversely proportional to the frequency sweep range, by:

$$\Lambda = \frac{c}{\Delta\nu} \quad (17)$$

Eq.(16) becomes:

$$I(OPD(t), t) = I_0 \left[ 1 + V \cdot \cos \left( 2\pi \left( \frac{n \cdot OPD(t)}{\Lambda} \cdot \frac{t}{\Delta t} + \frac{n \cdot OPD(t)}{\lambda} \right) \right) \right] \quad (18)$$

It is clear that the value of the OPD can contribute to the detected signal both in terms of the optical wavelength  $\lambda$  and/or in terms of the synthetic wavelength  $\Lambda$  generated by the frequency sweep.

To analyse Eq.(18) we consider two simple cases: 1 - no frequency sweep and, 2 - static OPD. In the first case,  $\Delta v = 0$ ,  $\Lambda = \infty$  and Eq.(18) becomes:

$$I(OPD(t), t; \Delta v = 0) = I_0 \left[ 1 + V \cdot \cos \left( 2\pi \left( \frac{n \cdot OPD(t)}{\lambda} \right) \right) \right] \quad (19)$$

In this case, as the OPD changes in time, detector will sense a complete fringe for every  $\lambda/n$  variation in the OPD. This is the particular case of relative single wavelength interferometry.

In the second case, the OPD is constant during the frequency sweep, and Eq.(18) becomes:

$$I(OPD_0, t) = I_0 \left[ 1 + V \cdot \cos \left( 2\pi \left( \frac{n \cdot OPD_0}{\Lambda} \cdot \frac{t}{\Delta t} + \frac{n \cdot OPD_0}{\lambda} \right) \right) \right] \quad (20)$$

The second term of the cosine argument represents only a constant phase and, as the frequency sweeps, the detector will sense a number of synthetic fringes that correspond to the absolute value of the OPD in  $\Lambda/n$  units. This is the ideal baseline for FSI. The absolute value of a distance  $L$  (half the OPD) in a FSI Michelson interferometer (Fig. 1) is obtained from the detected signal during the sweep interval  $\Delta t$  by:

$$L = \frac{N}{2} \cdot \frac{\Lambda}{n} \quad (21)$$

where  $N$  is the number of synthetic fringes detected from the beginning to the end of the sweep (integer and fractional part).

However, a third more realist case must be considered when the two phenomena, frequency sweep and OPD variation, occur simultaneously. In this case, the contribution of the two phenomena is mixed up in the detected signal, and it is impossible to distinguish between the individual contributions from each one of them.

It must be emphasised that, for a typical ADM application, the absolute value of the length to be measured is always several orders of magnitude larger than the length variations during the sweep (hereafter designated by drift). When drift occurs during the frequency sweep, two different types of fringes will be generated:

1. synthetic wavelength fringes ( $\Lambda$ -fringes) due to frequency sweep;
2. optical wavelength fringes ( $\lambda$ -fringes) due to changes in distance caused by drift.

As it is impossible to distinguish these two types of fringes,  $\lambda$ -fringes will be misinterpreted as  $\Lambda$ -fringes, causing the noise error (drift) to be multiplied by a large amplification factor  $\Lambda/\lambda$ .

The effect of drift in the FSI measurement can be separated in two different cases. If the drift signal (movement direction) changes during the sweep duration (usually less than a second), we are in presence of a low amplitude drift typically with frequencies identical or higher than the measurement frequency. In this case, the small amplitudes will have a smaller influence and, if it is not negligible, due to its random nature, it can be reduced applying a mean filter to the measurements.

In a second case, if the drift signal does not change during the sweep, we are in the presence of a low frequency drift with considerable amplitude even within the small time duration of the sweep (bearing in mind that this effect will be amplified). In this case, the drift must be compensated. Independent relative metrology and control can always reduce the drift by measuring it and by actuating a delay line during the sweep, or take it into account in the distance calculations. Nevertheless, this option would increase the complexity of the sensor, jeopardizing one of the FSI advantages.

If the OPD is written as:

$$OPD(t) = 2 \cdot (L_0 + l(t)) \quad (22)$$

where  $L_0$  is the distance at the beginning of the frequency sweep and  $l(t)$  the relative distance variation in time, the drift (the factor of 2 is due to the round trip in the measuring arm of the interferometer), the detected intensity will be given by:

$$I(t) = I_0 \left[ 1 + V \cdot \cos \left( 2\pi \left( \frac{n \cdot 2 \cdot (L_0 + l(t))}{\Lambda} \cdot \frac{t}{\Delta t} + \frac{n \cdot 2 \cdot (L_0 + l(t))}{\lambda} \right) \right) \right] \quad (23)$$

The absolute distance measurement is based on the fringe counting from the beginning ( $t = 0$ ) to the end of the frequency sweep ( $t = \Delta t$ ). As long as the drift signal does not change during the sweep, it is not important to know how the fringes change in time but only to measure the total number of detected fringes from the start to the end of the sweep. For this purpose, we need only to describe the drift as:

$$l(t) = \begin{cases} 0 & \text{for } t = 0 \\ \Delta L & \text{for } t = \Delta t \end{cases} \quad (24)$$

From Eq.(23) and (24) we can determine the number of detected fringes by subtracting the initial from the final phase of the cosine argument (divided by  $2\pi$ ):

$$N = \left[ \frac{n \cdot 2 \cdot (L_0 + \Delta L)}{\Lambda} + \frac{n \cdot 2 \cdot (L_0 + \Delta L)}{\lambda} \right]_{t=\Delta t} - \left[ \frac{n \cdot 2 \cdot (L_0)}{\lambda} \right]_{t=0} \quad (25)$$

resulting in:

$$N = \frac{2 \cdot n}{\Lambda} \left( L_0 + \left( 1 + \frac{\Lambda}{\lambda} \right) \cdot \Delta L \right) \quad (26)$$

From Eq.(26) it is possible to determine the absolute distance at the end of the sweep:

$$L = N \frac{\Lambda}{2 \cdot n} - \frac{\Lambda}{\lambda} \cdot \Delta L \quad (27)$$

Comparing Eq.(27) and (21), it is clear that the drift introduces an error equal to the drift amplitude multiplied by an amplification factor given by the ratio between synthetic and optical wavelengths.

When drift is neglected, compensation is not performed and the second term of Eq.(27) corresponds to an additional error in the measurement. As an example, for a sweep range of 50 GHz ( $\Lambda \approx 6$  mm) in the visible ( $\lambda \approx 600$  nm), the amplification factor is  $\approx 10\,000$ ; a drift with an amplitude of 1 nm would introduce an error of 10  $\mu\text{m}$  in the measurement. Depending on requirements, such an error may or may not be negligible.

It is thus very important to work with the lowest possible amplification factor either by using a longer laser wavelength (near IR) or a smaller synthetic wavelength by increasing the sweep range (without increasing the sweep duration otherwise the drift would also increase). Nevertheless, special care must be taken when the synthetic wavelength decreases, because the number of synthetic fringes will increase (which may also increase the complexity of the electronics counting subsystem).

### 3.1 Synthetic wavelength measurement and distance calculation

To calculate the absolute distance it is necessary to know not only the number of fringes but also the value of the synthetic wavelength, by measuring the value of the frequency sweep range.

In the experimental examples that will be discussed, the frequency sweep range measurement subsystem is based on a Fabry-Pérot interferometer (FP). The sweep range measurement is obtained by multiplying the FP free spectral range (FSR) by the number of resonances detected while the laser frequency sweeps. The beginning of the sweep is determined by a particular cavity mode; the laser sweeps from mode to mode and resonances, separated by the FP FSR, are detected and counted (Fig. 2). The frequency sweep range is then given by:

$$\Delta\nu = r \cdot FSR \quad (28)$$

where  $r$  is the number of detected FSR (number of resonances minus one).

From Eq.(17) and (28) we obtain:

$$\Lambda = \frac{c}{r \cdot FSR} \quad (29)$$

The measured length is, considering Eq.(27) and (29), given by:

$$L = N \frac{c}{2 \cdot n \cdot r \cdot FSR} - \frac{c}{\lambda \cdot r \cdot FSR} \cdot \Delta L \quad (30)$$

Without drift, or when its influence is smaller than the required uncertainty, we are in a Static Mode and the absolute distance can be calculated simply from:

$$L = N \frac{c}{2 \cdot n \cdot r \cdot FSR} \quad (31)$$

where  $c$ ,  $n$  and  $FSR$  are constants and  $N$  and  $r$  are the variables to be measured.

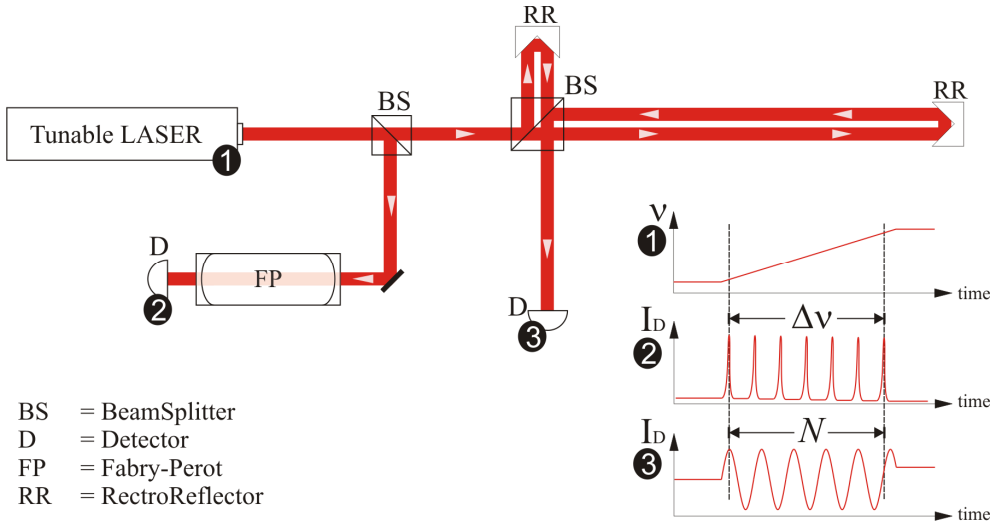


Fig. 2. FSI Setup. While the laser sweeps the frequency, the interferometer detector acquires the fringes and a FP measures the sweep range by counting the resonances of the cavity.

### 3.2 Drift compensation method for dynamic mode measurements

When it is not realistic to assume that the distance under measurement is constant during the sweep, the drift effect on the measurement has to be compensated. If we can assume that during the sweep there is no acceleration, i.e. the drift speed is constant, it is possible to remove the drift influence from two consecutive measurements with different values of sweep duration or/and sweep range. For a constant drift speed  $S_{drift}$ , the drift amplitude is described by:

$$\Delta L = S_{drift} \cdot \Delta t \quad (32)$$

and Eq.(27) can be written as:

$$L = N \frac{\Lambda}{2 \cdot n} - \text{sign}(\Delta v) \cdot \frac{\Lambda}{\lambda} \cdot S_{drift} \cdot \Delta t \quad (33)$$

where  $\text{sign}(\Delta v)$  is the sign of the frequency variation (+ for an increasing and - for a decreasing frequency sweep). For two consecutive measurements, 1 and 2, the following system of equations holds:

$$\begin{cases} L_1 = N_1 \frac{\Lambda_1}{2 \cdot n} - \text{sign}(\Delta v_1) \cdot \frac{\Lambda_1}{\lambda} \cdot S_{drift} \cdot \Delta t_1 \\ L_2 = N_2 \frac{\Lambda_2}{2 \cdot n} - \text{sign}(\Delta v_2) \cdot \frac{\Lambda_2}{\lambda} \cdot S_{drift} \cdot \Delta t_2 \\ L_2 = L_1 + S_{drift} \cdot \Delta t_3 \end{cases} \quad (34)$$

where  $\Delta t_3$  represents the time interval between the end of the first and the end of the second measurements. Drift speed and the corrected absolute distance for the second measurement are recovered by inverting Eq.(34) and using Eq.(29):

$$\begin{cases} S_{drift} = \frac{\lambda}{2 \cdot n} \cdot \frac{\left( \frac{N_2}{r_2} - \frac{N_1}{r_1} \right)}{\left( \Delta t_3 \cdot \frac{FSR}{c} \cdot \lambda \right) + \left( \text{sign}(\Delta v_2) \frac{\Delta t_2}{r_2} - \text{sign}(\Delta v_1) \frac{\Delta t_1}{r_1} \right)} \\ L_2 = \frac{c}{r_2 \cdot FSR} \cdot \left( \frac{N_2}{2 \cdot n} - \text{sign}(\Delta v_2) \cdot S_{drift} \cdot \frac{\Delta t_2}{\lambda} \right) \end{cases} \quad (35)$$

These equations are exact as long as drift speed is constant for the time span of the two measurements. Note that the frequency of the dynamic mode measurement can still be the same as it would be in the static mode single measurement as the first measurement can always be the same as the second measurement of the previous pair of sweeps. Thus, every new measurement enables another dynamic mode measurement (in this case, the characteristics of the two sweeps would swap from measurement to measurement).

The most simple implementation of this compensation method is to use two consecutive measurements with different signs on the sweep range ( $\Delta v$ ) and equal durations, corresponding to a symmetrical triangular shaped frequency sweep ( $\Delta t_2 = \Delta t_1 = \Delta t$  and  $\Delta v_2 = -\Delta v_1 = \Delta v \Rightarrow r_2 = r_1 = r$ ). In this case, we obtain:

$$\begin{cases} S_{drift} = \frac{\lambda}{2 \cdot n} \cdot \frac{N_2 - N_1}{\left( \Delta t_3 \cdot \frac{r \cdot FSR}{c} \cdot \lambda \right) - 2\Delta t} \\ L_2 = \frac{c}{r \cdot FSR} \cdot \left( \frac{N_2}{2 \cdot n} + S_{drift} \cdot \frac{\Delta t}{\lambda} \right) \end{cases} \quad (36)$$

With this approach, we are not only correcting the length measurement but also providing a measurement of the drift speed.

### 3.3 Measurement performances

FSI sensor performance does not depend on the stability of the absolute value of the frequency but on the uncertainty in the frequency sweep range: there is no need to calibrate the system for absolute frequencies.

The system final uncertainty has two major contributors:



- uncertainty in the measured number of fringes  $N$  (the synthetic fringe interpolation uncertainty);
- uncertainty in the sweep range measurement  $\Delta v$ . The later depends on:
  - uncertainties in the  $FSR$  value and on
  - determination of the number of detected  $FSR$  ( $r$ ).

Without drift (static mode), the final uncertainty, corresponding to a coverage probability of approximately 95% (IOS, 1995), can be obtained from (31):

$$\delta L = \sqrt{\sum_i \left( \frac{\partial L(X_i)}{\partial X_i} \cdot \delta X_i \right)^2} \quad (37)$$

$$\delta L = \sqrt{\left( \frac{c}{2 \cdot n \cdot FSR \cdot r} \cdot \delta N \right)^2 + \left( \frac{c \cdot N}{2 \cdot n \cdot FSR^2 \cdot r} \cdot \delta FSR \right)^2 + \left( \frac{c \cdot N}{2 \cdot n \cdot FSR \cdot r^2} \cdot \delta r \right)^2 + \left( \frac{c \cdot N}{2 \cdot n^2 \cdot FSR \cdot r} \cdot \delta n \right)^2}$$

The uncertainty component in  $N$  is only the fringe phase uncertainty,  $\delta N$ , because the integer number of fringes can be considered to be measured exactly. In the determination of the integer part, if a fringe is missed or over counted, the effect is easily noticed and can be corrected using simple outlier removal procedures. This uncertainty does not change with range, and depends only on the value of the synthetic wavelength. This component is dominant for small distances because it does not increase with  $L$ .

The next two components, in  $r$  and in  $FSR$ , determine the uncertainty in the sweep range measurement and are related to the FP performances.

The uncertainty in  $r$  depends on FP finesse and on the capability of the signal processing to localize resonance maxima. In order to locate resonance maxima in time unambiguously, the maxima of the signal generated by the FP should be clearly discriminated. The higher the finesse, the easier it is to locate accurately each resonance maxima in time and, therefore, reduce derived sensor errors.

The value of the  $FSR$  uncertainty is determined by the FP calibration and stability. The stability depends on length and optical variations induced by temperature and misalignments. For high resolution (low  $FSR$  uncertainty), thermal stabilization and low-thermal expansion materials may be required.

In contrast to the uncertainty in the synthetic fringe interpolation, the uncertainty components related to the sweep range measurement increase with range. These components become dominant for large distances.

The refractive index of air contributes to the measurement uncertainty due to the instabilities in the optical medium caused by the variations of the air temperature, pressure

and humidity. The value of the refractive index of air can be determined by the Edlén equation (Stone & Zimmerman, 2004). The different parameters contribute differently to the index uncertainty. As an example, for normal laboratory conditions ( $p = 100,4$  kPa;  $T = 20$  °C;  $RH = 50\%$ ), the contribution given by a change in the humidity of 50% is equivalent to a change of 0,2 kPa or 0,5 °C. The influence of the humidity is thus irrelevant compared to the influence of pressure and temperature. In order to neglect the contribution of the laboratory environment the refractive index contribution should be kept at least around  $10^{-6}$ . This can be achieved by measuring the pressure with an uncertainty around  $\pm 0,2$  kPa ( $\pm 2$  mbar) and by measuring (and controlling) the temperature within  $\pm 0,5$  °C. Note that for a space application this factor is null.

Fig. 3 shows an example of the system performances for these three components as a function of distance, considering typical parameters (presented in Table 1) used in the experimental results to shown later.

The value of the PF FSR should be selected in order to optimize its contribution to the final uncertainty. For the same sweep range, the number of resonances decreases when the FSR increases. Thus, if the  $\delta FSR$  contribution is larger than the  $\delta r$  contribution, the value of the FSR should increase and vice-versa. It must be noted that a low FSR is limited by the maximum available etalon length (the length of the FP increases when the FSR decreases) and by the FSR uncertainty (as the FSR becomes smaller, the number of resonances for the same sweep range increases and the contribution of the FSR uncertainty increases as well).

The previous analysis corresponds to a driftless static mode. In the dynamic mode with a constant drift speed, the same analysis can be performed starting with Eq. (35) or (36), leading in this case to a more complex solution. Nevertheless, the conclusions are basically the same.

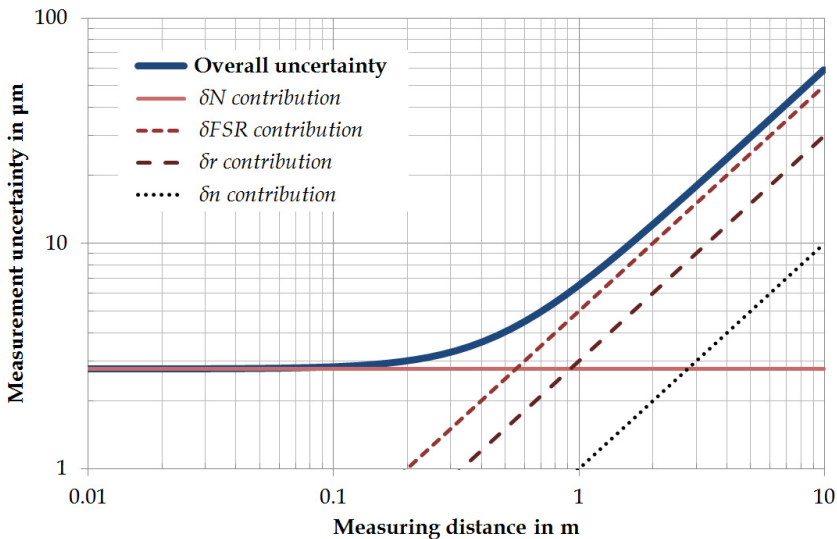


Fig. 3. Example of the FSI model, considering typical parameters presented in Table 1. ( $\delta N = 1/360$  of a fringe,  $\delta FSR = 5 \cdot 10^{-6}$ ,  $\delta r = 3 \cdot 10^{-6}$ ,  $\delta n = 1 \cdot 10^{-6}$ ).

<i>Parameter</i>	<i>Value</i>	<i>Uncertainty</i>
Frequency sweep range	$\Delta\nu = 150 \text{ GHz}$	$\delta\Delta\nu = 750 \text{ kHz}$ $5.10^{-6} \text{ Hz/Hz}$
Synthetic wavelength	$\Lambda = 2 \text{ mm}$	$\delta\Lambda = 10 \text{ nm}$ $5.10^{-6} \text{ m/m}$
Synthetic fringes		$\delta N = 1/360$
FP FSR	$FSR = 1 \text{ GHz}$	$\delta FSR = 5 \text{ kHz}$ $5.10^{-6} \text{ Hz/Hz}$
Number of FSR	$r = 150$	$\delta r = 3.10^{-6}$
Refractive index	$n = 1,00027$	$\delta n = 1.10^{-6}$

Table 1. FSI parameters used in the simulation presented in Fig. 3.

#### 4. Dual FSI concept for large range measurements

As mentioned before, as we increase the measurement range, accuracy and complexity are typically linked together. This is due to the fact that in FSI the measurement uncertainty increases with distance, as a consequence of the uncertainty propagation of the synthetic wavelength measurement (directly related to the FP performance). In order to achieve the required uncertainty one would have to increase the FP stability (thermal and mechanical) and improve the calibration and resonance detection characteristics (i.e. using Pound-Drever-Hall technique to lock the laser into FP resonances). As a consequence, the expected impact on sensor complexity would be critical. For large ranges, this increase in uncertainty can be considered the main drawback of the technique.

To overcome this problem, maintaining the low complexity associated to short distances, the measurement process for larger ranges can be reduced to the close range case, by limiting the OPD in the interferometer. This can be achieved by increasing the reference arm with a long reference fibre and introducing the concept of dual FSI mode, where an ancillary interferometer is used to measure (calibrate) continuously the fibre length (Cabral et al, 2009).

By adding a long reference fibre in the reference arm, the measured OPD is reduced to the real distance minus half fibre optical path length. Using an additional interferometer, the fibre is calibrated with the same accuracy but with lower requirements. This assumes that the fibre length is constant, during a certain time interval, and therefore the calibration will be the result of averaging for that period (uncertainty is reduced by a factor proportional to the square root of the number of measurements).

In this dual FSI approach, the final uncertainty is the sum of two components:

- uncertainty in the measurement of the reduced OPD (twice the absolute distance minus Reference Fibre length);
- calibration uncertainty of the Reference Fibre Optical Path Length (OPL);

In addition, during the calibration period, we must guarantee that the change in fibre OPL due to thermal variation is negligible compared with the required calibration uncertainty. It must be noted that the change in fibre length due to thermal variation is not the only limiting factor to the number of measurements that can be averaged for fibre calibration. Beyond a certain reduction factor, i.e. after a certain averaging period, the uncertainty

component related to the FP FSR calibration uncertainty becomes dominant and there is no benefit in increasing the calibration period.

Besides the measurement dispersion, there are other uncertainty components that can influence the final measurement accuracy. The resolution in the fringe phase and resonance position measurements is currently much smaller than the measurement dispersion, and thus its contribution is not significant. The uncertainty in the calibration of the FP FSR is the most important one for large distance measurements in a single FSI configuration but, in the dual approach, it is possible to achieve a very small contribution considering the value of the dispersion component.

In order to maintain low hardware complexity, the optical setup for the implementation of this Dual FSI concept makes use of the same laser and FP for the two interferometers, and also some common beam splitters (BS). For that, a Mach-Zehnder configuration was selected, and is presented in Fig. 4, where the two interferometers are shown separately (note that BS1, BS2 and the reference fibre are common to both interferometers).

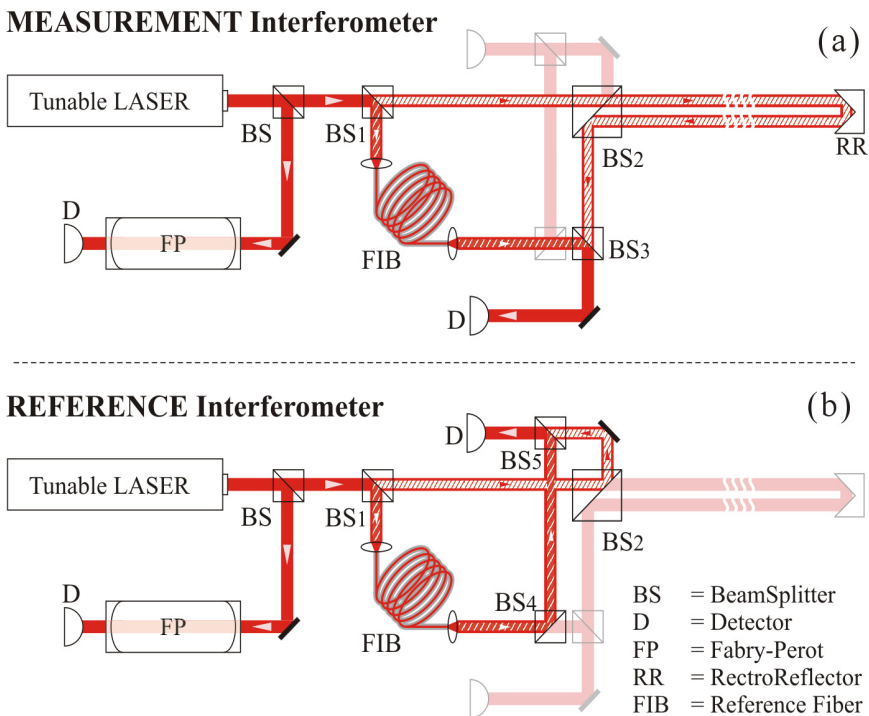


Fig. 4. Dual FSI interferometers setup, measurement (a) and reference (b).

The measurement interferometer evaluates the OPD between the OPL of twice the distance in measurement (due to the round trip) and the OPL added by the reference fibre, as shown in Fig. 4(a). The reference interferometer, shown in the Fig. 4(b), will measure the OPL added by the reference fibre that cause the reduction of the OPD in the measurement interferometer.

## 5. Experimental implementation of FSI based sensors

In order to demonstrate the capabilities of FSI based sensors, two implementations are described in this section, one for FSI and another for Dual FSI, both tailored for space applications (i.e. long distances, no atmosphere refraction errors are considered, specific operational requirements associated to measurement rate and accuracy).

The first absolute distance sensor, based in FSI, was designed for a space mission comprising a multiple aperture optical telescope (coherent array), to measure the OPD between two sub-telescopes and the combining-telescope with an uncertainty at the 10  $\mu\text{m}$  level at a 10 Hz measurement rate with the lowest possible technical complexity.

The second sensor, based in Dual FSI, was defined in view of a coronagraph mission composed of two spacecraft in a free-flying formation. The requirements on the absolute distance metrology sensor prototype were a measurement uncertainty below the 100  $\mu\text{m}$  level for a distance around 50 m at a 10 Hz measurement rate.

In both sensors, the Laser and the FP are the key items. Complex tasks normally implemented in real time electronics were assigned to data processing, thus moving system complexity to the software area.

A drift compensation model using consecutive measurements was implemented and validated experimentally. Fig. 5 shows the principle behind the compensation technique using two consecutive measurements with different signs on the sweep range ( $\Delta v$ ) and equal durations, corresponding to a symmetrical triangular shaped frequency sweep ( $\Delta t_2 = \Delta t_1 = \Delta t$  and  $\Delta v_2 = -\Delta v_1 = \Delta v \Rightarrow r_2 = r_1 = r$ ).

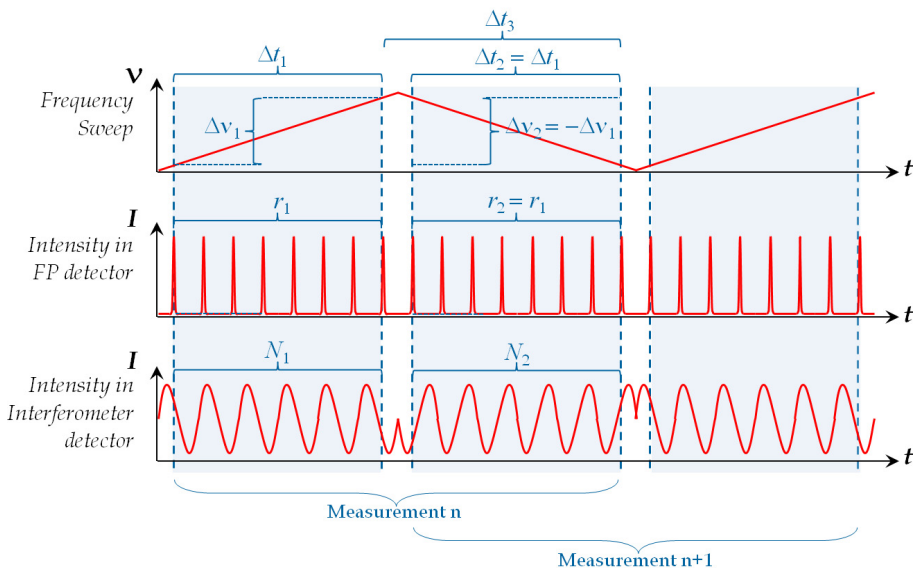


Fig. 5. Principle of the Drift compensation model using two consecutive measurements with different signs on the sweep range and equal durations corresponding to a symmetrical triangular shaped frequency sweep.

Note that the frequency of the dynamic mode measurement can still be the same as it would be in the static mode single measurement as the first measurement can always be the same as the second measurement of the previous pair of sweeps. Thus, every new measurement enables another dynamic mode measurement (in this case, the characteristics of the two sweeps would swap from measurement to measurement). With this approach, we are not only correcting the length measurement but also providing a measurement of the drift speed, using Eq.(36).

### 5.1 Tunable laser

FSI needs a tunable laser, the frequency of which changes continuously without mode hop. The tunable range should be able to cover several tens of GHz, and this range must be completely swept in a time frame not exceeding a tenth of second. Laser linewidth must be small in order to generate a coherence length of several tens of metre. It is also important to have a compact, robust and with low complexity laser.

Although several options exist to provide wavelength/frequency tuning, the best solution to cope with the requirements is an External Cavity Diode Laser (ECDL) (Amann & Buus, 1998). Traditional diode-laser technology provides high reliability, high electrical efficiency, and a wide range of available wavelengths. By using an external cavity built around a diode laser, it is possible to achieve single-mode operation with narrow linewidth and precise frequency tuning. The extended cavity leads to a spectral narrowing of the laser output to below 1 MHz. The coherence length can thus be in the order of a few hundred metres.

### 5.2 Fabry-Pérot interferometer

In FSI implementation described, the frequency sweep range measurement subsystem is based on a FP interferometer (Vaughan, 1989). As the sweep range measurement is obtained by multiplying the FP FSR by the number of resonances detected while the laser frequency sweeps, the FP interferometer is a critical component of the sensor.

A FP is a linear resonator which consists of two highly reflecting mirrors forming a standing-wave cavity. In this case, we exploit the fact that the transmission through such a cavity exhibits sharp resonances, making it perfect as a frequency reference cavity.

Two basic types of FP cavity exist: planar and confocal. The confocal is more appropriate for the current application as it has inherent higher finesse, is much less sensitive to mirror misalignment (because it is not critical to maintain mirror parallelism), and for the same FSR the length of the confocal cavity is half the length of the planar cavity. When a confocal FP is illuminated by a monochromatic beam close to the axis, a multiple beam interference pattern is produced near the centre of the interferometer. At precisely the confocal spacing, each mirror images the other back upon itself so that a paraxial ray is re-entrant after four traversals of the cavity. This means that the transmitted spectrum is reproduced with every quarter wavelength change in the mirror separation. Consequently, the FSR is given by:

$$FSR = \frac{c}{4 \cdot n \cdot d} \quad (38)$$

where  $c$  is the speed of light,  $n$  is the refractive index of the optical medium between the mirrors, and  $d$  is the distance between the spherical mirrors (whose radius of curvature is equal to the spacing between them).

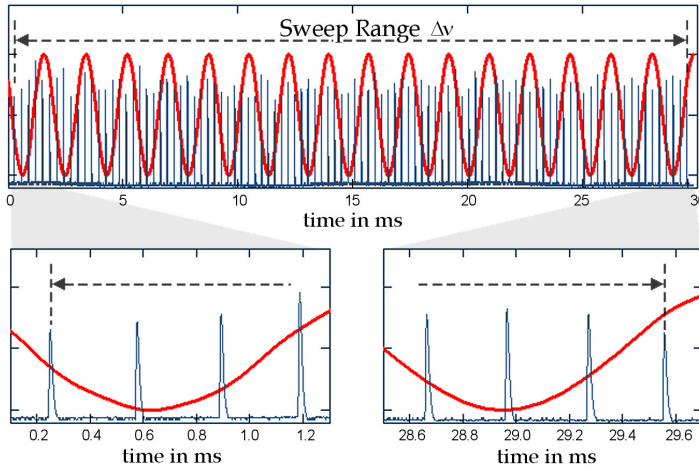


Fig. 6. Typical FSI range and, consecutively, the frequency sweep range value and its uncertainty.

### 5.3 Data processing

The data acquisition of the FSI interferometer is based in a simple homodyne detection scheme. During the measurement, while the laser frequency is sweeping, signals from the detectors of the FP and interferometer (one or two depending if it is FSI or Dual FSI) are acquired simultaneously. As the maximum number of fringes for the required range is small, it is conceivable to acquire all the data for subsequent processing. After signal processing, we obtain the number of fringes per resonance  $N/r$ .

As the requirements on the sampling rate of the data acquisition subsystem increase, the system should be able to keep only the initial and final parts of the signal with high resolution and acquire the middle data at a lower sampling rate - system performances depend only on the accurate determination of the initial and final resonances and zero crossing positions (another possibility is to lock the laser frequency to the initial and final resonances, although this solution increases complexity).

To retrieve distance from the acquired signals using Eq. (31), the data processing chain must:

1. define the sweep range by choosing the initial and final frequency resonances;
2. determine the sweep range by measuring the number of Fabry-Pérot resonances within the sweep range;
3. measure the number of synthetic fringes (both integer and fractional part) within the sweep range;
4. calculate the number of fringes per resonance.

Data processing can be more efficient if all the available data is used. Instead of measuring only  $N$  and  $r$  to obtain  $N/r$ , we can measure the number of fringes  $N_i$  for each resonance  $r_i$  (where  $i$  represent the sequence of resonances) and determine  $N/r$  using a linear regression, thus reducing the uncertainty in the measurement, and making the procedure more robust to errors (we are making measurements at several tens of resonances instead of only two, the initial and the final).

The control, acquisition and data processing applications were all custom made, implemented in Labview™ from National Instruments. The system has the capability to do signal acquisition up to 5 MS/s (16 bits) and the output (for laser control) up to 4 MS/s (16 bits).

**5.4 FSI sensor implementation and results**

The first example of an FSI based sensor was designed to measure the OPD between two sub-telescopes of a multiple aperture optical telescope (Cabral & Rebordão, 2007). The implemented setup is sketched in Fig. 7. In the case of an absolute OPD measurement the result may be either positive or negative. The sign of the absolute OPD is thus required and must be measured. To overcome the ambiguity, a calibrated offset distance is created in arm A ( $A' = A + \text{Offset}$ ) of the interferometer (Fig. 7) ensuring that  $A' - B > 0$ , always, with the cost of increasing the maximum range (the Offset must be larger than half of the measurement range). The real distance difference is obtained by subtracting the calibrated Offset value.

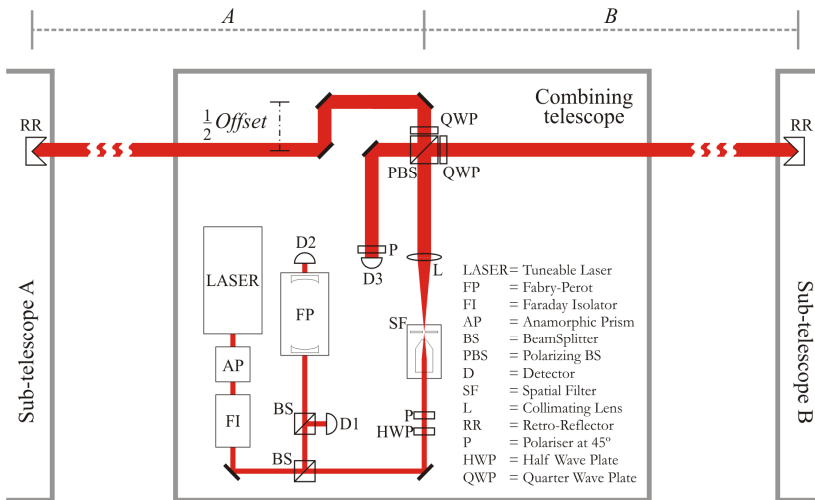


Fig. 7. Setup of the implemented FSI absolute sensor.

A Faraday isolator (FI) prevents feedback from the interferometer components back to the ECDL laser, and an anamorphic prism pair (AP) circularizes the elliptical cross section of the beam. The beam is then split by the first beam splitter (BS): one part is reflected to the sweep range & power monitoring measurement subsystems and the remainder is transmitted to the interferometer. A second BS splits the reflected beam into the power monitoring detector (D1) (reflected part) and the FP Interferometer that includes D2 (transmitted part). The beam



transmitted to the interferometer is filtered by the spatial filter (SF) and collimated by lens (L); before entering the interferometer, it goes through a half wave plate (HWP) and a 45° oriented polarizer (P). These two elements enable the original vertical polarized light to be rotated by 45°. After the interferometer beam splitter, two quarter wave plates (QWP) transform the reflected vertical and transmitted horizontal polarized into circular polarized light that, after reflection in the retro-reflectors (RR) and another passage through the QWP, are redirected towards the detector (D3). Before the detector, a 45° polarizer (P) enables interferences.

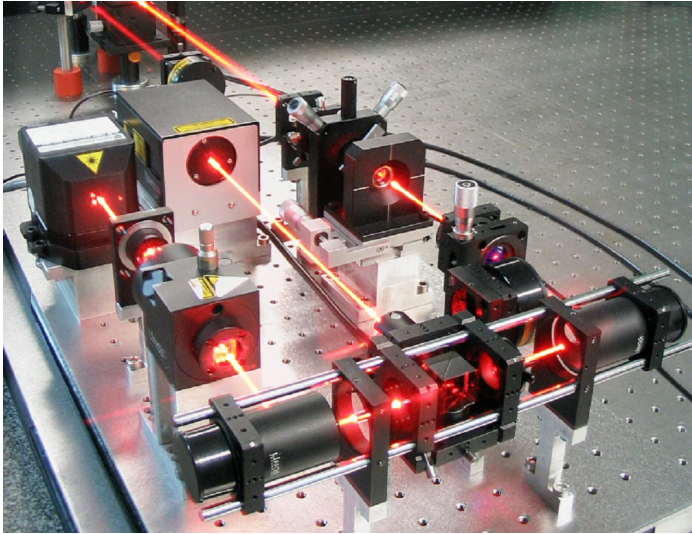


Fig. 8. FSI prototype implementation of the setup described in Fig. 6.

Fig. 8 shows a picture of the implemented prototype. The tunable laser is an ECDL capable of a mode-hop-free sweep range up to 150 GHz (TBL-7000 from NewFocus in a Littman-Metcalf mounting at 633 nm), corresponding to a synthetic wavelength down to 2 mm. The FP has a free spectral range of 1 GHz (TOPTICA high resolution, temperature stabilized, confocal Fabry-Pérot with a real finesse up to 1000). The FP was calibrated using the sensor hardware by means of two absolute measurements and a relative calibrated distance measurement between the two absolute positions, a kind of self-calibration (Cabral & Rebordão, 2006). This is a simple procedure and thus highly convenient to be implemented in space applications.

The graph in Fig. 9 shows the typical results obtained with this sensor. Each point corresponds to 200 consecutive measurements (with sweep duration of 100 ms) and a  $2\sigma$  m for distances up to 1 m, increasing with distance as expected. Results confirm model predictions (continuous line) for  $\delta N = 1/500$ ,  $\delta FSR = 6$  kHz, and  $\delta r = 1/3000$  (consequently  $\delta \Lambda = 14$  nm). Measurement uncertainty is well below the 10  $\mu$ m level.

Similar tests were performed for the dynamic mode that confirmed the efficiency of the compensation model, even in the presence of a non-constant drift speed (Cabral & Rebordão, 2007).

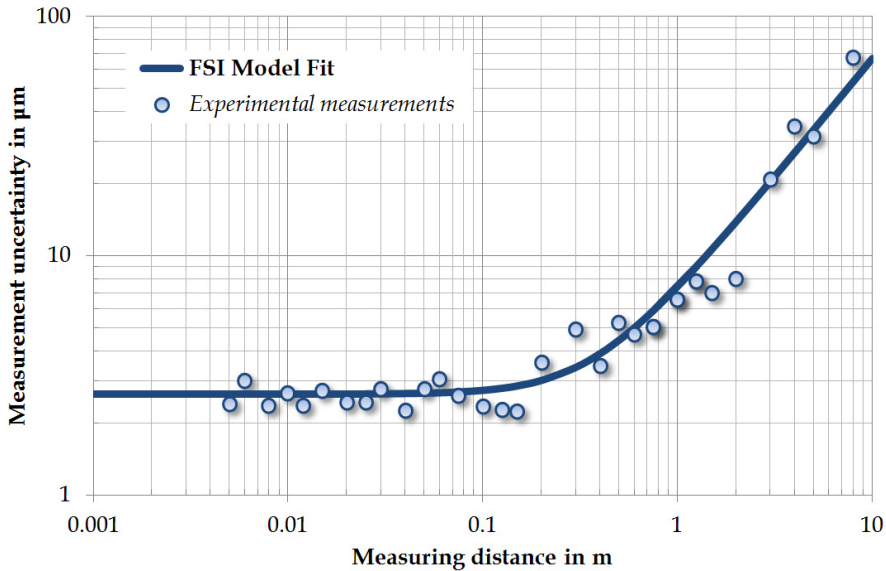


Fig. 9. FSI experimental measurements  $2\sigma$  error in static a mode.

### 5.5 Dual FSI sensor implementation and results

The second example of an FSI based sensor makes use of the Dual FSI concept. The prototype was designed to measure at a nominal distance of 50 m, within a 10 m interval (Cabral et al., 2010). As mentioned previously, for such a distance, in a traditional FSI implementation (single interferometer mode), the uncertainty is dominated by the two uncertainty components related to the sweep range measurement and, to achieve accuracies below 100  $\mu\text{m}$  it would be necessary to increase significantly the complexity of the FP subsystem. The Dual FSI approach overcomes this limitation without a major increase in the complexity, simply by adding a long reference fibre in the reference arm of a measurement interferometer and using an additional (reference) interferometer to calibrate the fibre.

In order to provide the sensor with assembly flexibility, the prototype implementation comprises three subunits, mechanically decoupled, with an optical connection implemented by fibres:

- Laser & Detection – Unit that contains the ECDL and the detectors for the two interferometers and the FP.
- FSI Head – Unit that includes the FP and the FSI interferometer(s) beam splitter (BS1 in Fig. 4).
- Optical Head – Optics that form the two interferometer arms.

Fig. 10 shows a picture of the Dual mode FSI sensor breadboard where is possible to identify the Laser & Detection, the FSI Head, the Optical Head unit and the long reference fibre housing.

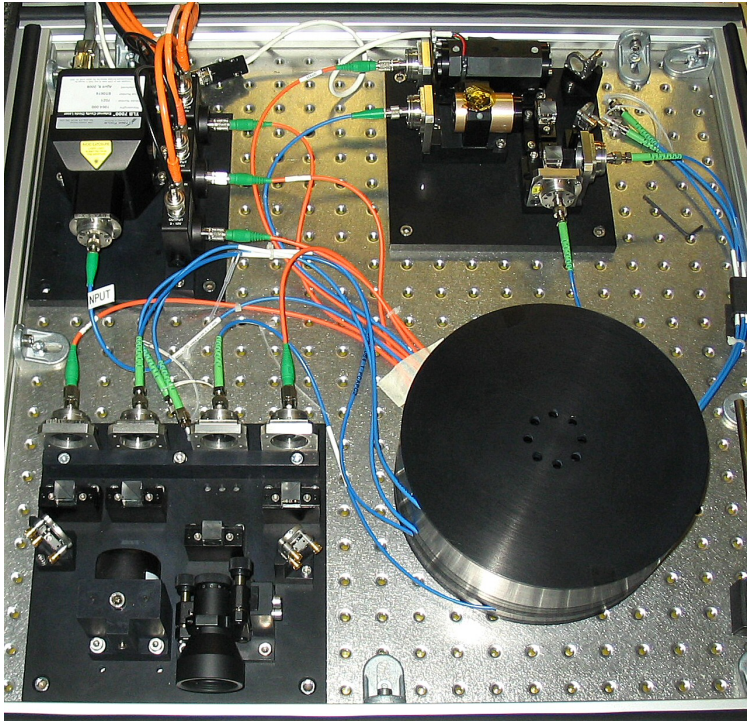


Fig. 10. Dual FSI sensor breadboard, showing Laser & Detection unit (top left), the FSI Head (top right), the Optical Head (bottom left) and the long reference fibre housing (bottom right).

In the Laser & Detection unit, the laser is the StableWave™ Model TLB-7021 by NewFocus, based on a Littman-Metcalf design (1024 nm). This ECDL is fibre pigtailed, and can achieve a maximum 85 GHz mode-hop-free frequency sweep, resulting in a minimum synthetic wavelength of 3.5 mm. In terms of sweep speed, it is possible to perform the complete sweep range in 50 ms, using a triangular shaped modulation, without noticing any major change in the laser response. The laser has a good stability in the short-term ( $\mu\text{s}$ ), shorter than 50 kHz, but a poor stability performance at the medium-term (ms), having a direct impact in the sensor performance (disturbances caused by acoustically excited mechanical resonances in the laser cavity).

The FSI Head starts with a Faraday Isolator to prevent any residual back reflection into the laser as the ECDL behaviour is highly sensitive to feedback. Next, the light is split to the interferometers and the FP. As mentioned before, the FP has a critical impact in the system performances. An air spaced confocal etalon was manufactured (by IC Optical Systems Ltd) in Zerodur with a cavity length of 50 mm, corresponding to a FSR of 1.5 GHz. In terms of Finesse, the selected high reflectivity coating produces a very thin resonance peaks, experimentally measured to be, at least, 4700. The measured level of thermal stability (approximately 10 mK) corresponds to a FSR stability of a few tens of Hz, two orders of magnitude lower than the requirement in the uncertainty of the FSR.

After the beam splitter that directs half the light to the FP, the light is split again towards a small connection fibre and the long reference fibre, both already part of the interferometers arms. Both fibres are PANDA Polarization Maintaining (PM) with FC/APC connectors (low return loss). All the fibres are located in the fibre housing, shown in Fig. 11, to ensure mechanical stability and increase (by inertia) the short term thermal stability.

The length of the long reference fibre, used in the Dual FSI interferometers, must be twice the distance we want to subtract to the measuring range (while the light in the measurement path performs a round trip in the reference fibre it only travels in one direction). A 71 m fibre was selected, allowing an OPL subtraction of approximately 102 m (corresponding to the difference between the reference fibre length and the measurement arm fibre of 1 m multiplied by the refractive index of the fused silica). The setup of the Optical Head is slightly different from the one illustrated in Fig. 2, as in the reference interferometer an additional retro-reflector (RR<sub>REF</sub>) allows the definition of the point from where the absolute measurement is referenced. Fig. 11 shows the optical setup of the implemented measurement (a) and reference interferometer (b). The position of the BS was chosen in order to make the measured OPD in the reference interferometer equal to the difference between the two fibres ( $F_{REF} - F_{MEAS}$ ).

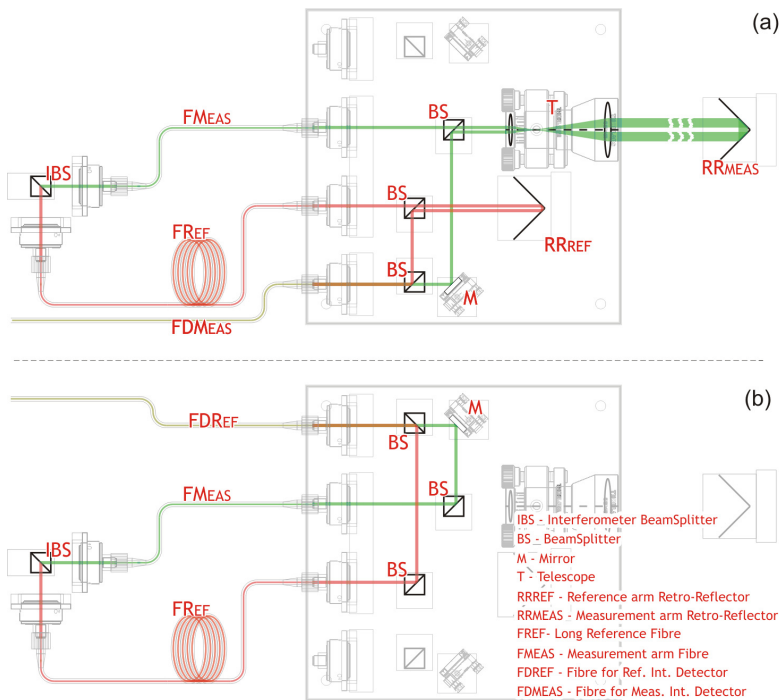


Fig. 11. The dual FSI measurement (a) and reference (b) interferometer.

The performances of the sensor result on the added contribution of the uncertainty in the reduced OPD measurement (by the measurement interferometer) and the long reference fibre calibration (by the reference interferometer).

The graph in Fig. 12 shows the results obtained with the contribution of the two mentioned uncertainty components. Each point corresponds to the  $2\sigma$  dispersion (95% confidence interval) for 300 measurements. As shown, the contribution of the reduced OPD measurement is lower than  $10\ \mu\text{m}$  for the first few meters, and smaller than  $30\ \mu\text{m}$  for an absolute distance of 10 m.

To evaluate the component related with the calibration of the reference fiber, 2500 sequential measurements were performed, each with a duration of 100 ms. The measurement dispersion at  $2\sigma$  was  $554\ \mu\text{m}$  for a mean value of  $101\ 805\ 363\ \mu\text{m}$ . The final uncertainty in the calibration of the fiber OPL is a function of the number of measurements that contribute to the calculated average length (and also the contribution of the FP FSR calibration uncertainty). The longer the duration of the calibration, the larger will be the decrease (by  $\sqrt{N}$ ) of the uncertainty value resulting from the measurement dispersion. With a calibration period of 300 s (during which temperature was stable enough to consider a static OPL in the fibre), the uncertainty for the 101.8 m fibre OPL  $31.2\ \mu\text{m}$  (note that the contribution to the measured distance will be half of this value).

Fig.12 shows the uncertainty for the current dual FSI configuration using a 71 m fused silica fiber (101.8 m OPL) that allows a measurement range from 51 m to 61 m with accuracy smaller than  $40\ \mu\text{m}$ .

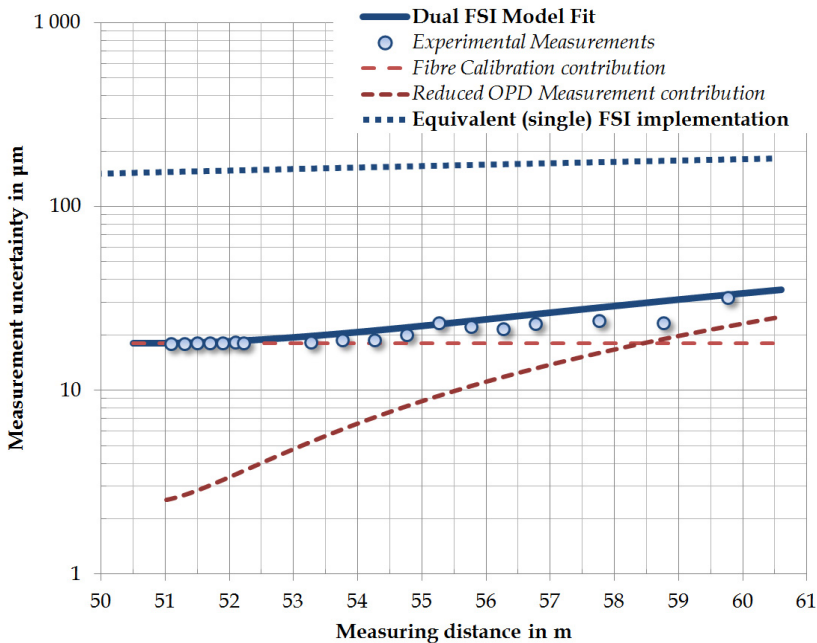


Fig. 12. Uncertainty for the current dual FSI configuration using a 71 m fused silica fibre (101.8 m OPL).

As it can be seen in Fig. 12, the improvements resulting from the Dual FSI concept compared to the expected results in a single FSI implementation are clear. Although the concept is

limited by configuration to specific range intervals (determined by the length of the reference fiber), the enhancement in the measurement accuracy is about one order of magnitude.

## 6. Conclusions

This chapter showed that FSI is a particularly versatile technique for absolute dimensional metrology, allowing different approaches for different applications with different needs, both in terms of range and accuracy.

The concept has a simple model, experimentally validated, the results of which are presented in this work. There is an excellent match between measurements and computer simulations. The model proved to be sufficiently robust to allow the setup of the different parameters and implementation concepts for particular applications.

The most critical issue in the FSI technique is sensitivity to drift. Whilst in a laboratory with a controlled environments (temperature, pressure and vibrations) it is possible to achieve conditions where drift can be neglected, the same does not apply to the majority of the applications (like the example of space instrumentation). The case of FSI as a low complexity sensor requires self-sufficiency in terms of drift compensation. A drift compensation method using two consecutive measurements was also presented. In addition to the compensation of the error introduced by drift, the method also provides an efficient way to measure drift speed, using exclusively the same set of data.

In terms of FSI performance, the accuracy at small distances is determined by the capability to interpolate a synthetic fringe. To improve accuracy we must choose the smallest possible synthetic wavelength, corresponding to the highest possible mode hop free frequency sweep range. The sweep range and the measurement frequency (which defines the sweep duration) determine the required frequency sweep rate. Even in an ECDL (currently the fastest tunable laser), the sweep rate should not exceed a few thousands of GHz/s, otherwise sweep non-linearities will jeopardise sensor performance. Therefore, even if the laser is capable of a larger sweep, the measurement frequency will limit the minimum synthetic wavelength. A smaller synthetic wavelength also decreases the influence of the drift because it decreases the amplification factor.

For large distances, accuracy is determined by the uncertainty in the synthetic wavelength value, as this will be propagated through the large number of synthetic fringes. As the synthetic wavelength is measured by counting the number of FSR in a FP, its uncertainty is determined by the stability of the FSR and also by how accurately the resonance peaks are located. Whenever the range is limited to a region around a large distance, it is possible to achieve high accuracy at large distance using a Dual FSI sensor approach, maintaining the reduced complexity inherent to the FSI technique.

Frequency sweeping interferometry is therefore an essential method for absolute distance metrology, especially when complexity and robustness are critical drivers (a must in space metrology). The flexibility of its parameterisation is an important advantage of the technique, allowing the same device to be used for a wide variety of mission requirements.

## 7. References

- Amann, M.-C. & Buus, J. (1998). *Tunable Laser Diodes*, Artech House ISBN: 978-0890069639, 1998
- Cabral, A. & Rebordão, J. (2005). Absolute distance metrology with frequency sweeping interferometry, *Recent Developments in Traceable Dimensional Measurements III*, J. Decker, Gwo-Sheng Peng; eds. Proc. SPIE 5879, 195-204, 2005
- Cabral, A. & Rebordão, J. (2006). Calibration of the Fabry-Pérot free spectral range using a tunable laser in a Michelson interferometer, *Optical Engineering* 45, 1005, 2006
- Cabral, A. & Rebordão, J. (2007). Accuracy of frequency sweeping interferometry for absolute distance metrology", *Optical Engineering* 46(07), 073602, 2007
- Cabral, A. et al. (2009). Absolute distance metrology for long distances with dual frequency sweeping interferometry, *XIX IMEKO World Congress, Fundamental and Applied Metrology*, Lisbon, Portugal, 1942-1947, 2009
- Cabral, A. et al. (2010). Dual-frequency sweeping interferometry for absolute metrology of long distances *Optical Engineering* 49, 085601, August 2010
- Calvel, B. et al. (2004). High Precision Optical Metrology for DARWIN, *2nd Int. Symposium on Formation Flying Missions and Technologies*, Arlington, VA, USA, 14-16, 2004
- Coe, P. A. et al. (2004). Frequency scanning interferometry in ATLAS: remote, multiple, simultaneous and precise distance measurements in a hostile environment, *Measurement Science and Technology*, vol. 15, pp. 2175-2187, 2004
- Dändliker, R. et al. (1988). Two-wavelength laser interferometry using superheterodyne detection, *Optical Letters*. 13(5), 339-341, 1988
- Dändliker, R. et al. (1995). High-accuracy distance measurements with multiple-wavelength interferometry, *Optical Engineering* 34(8), 2407-2412, 1995
- Edwards, C. S. et al. (2000). Experimental verification of a swept-frequency laser interferometer for distance measurement, *NPL Report CBTLM 4*, 2000
- Hecht, J. (2001). Tuning in to tunable lasers, *Laser Focus World*, pp.121-125, 2001
- IOS (1995). Guide to the Expression of Uncertainty in Measurement, *International Organization for Standardization*, Geneva, Switzerland, 1995
- Kikuta H. et al. (1986). Distance measurement by the wavelength shift of laser diode light, *Applied Optics*, vol. 25(17), pp.2976-2986, 1986
- Kinder, Th. & Salewski, K. D. (2002). Absolute distance interferometer with grating-stabilized tunable diode laser at 633 nm, *Journal of Optics A*, 4, S364-S368, 2002
- Stone, J. et al (1999). Absolute interferometry with a 670-nm external cavity diode laser, *Applied Optics* vol.38(28), pp.5981-5994, 1999
- Swinkels, B. et al. (2005). Absolute distance metrology for space interferometers, *Recent Developments in Traceable Dimensional Measurements III*; Jennifer E. Decker, Gwo-Sheng Peng; eds. Proc. SPIE 5879, 216-222, 2005
- Thiel, J. et al. (1995). Interferometric measurement of absolute distances of up to 40 m, *Measurement* vol.16, pp.1-6, 1995
- Vaughan, J.M. (1989). *The Fabry-Pérot Interferometer: History, Theory, Practice and Applications*, Institute of Physics Publishing, ISBN: 978-0852741382, 1989
- Yang, H. J. et al. (2005). High-precision absolute distance and vibration measurement with frequency scanned interferometry, *Applied Optics*, 44(19), 3937-3944, 2005
- Stone, J. & Zimmerman, J. (2004). Engineering metrology toolbox, refractive index of air calculator NIST, <http://emtoolbox.nist.gov/Wavelength/Documentation.asp>, 2004

# Nano-Accuracy Surface Figure Metrology of Precision Optics

Shinan Qian and Peter Takacs  
*Brookhaven National Laboratory*  
USA

## 1. Introduction

Precision optical surface figure measurement is the most important aspect for optical systems development of telescopes, microscopes, cameras and imaging or focusing systems. Any optical surface slope error or surface figure error departure from the theoretical surface will produce beam deviations according to Snell's law (law of refraction) for transmission optics or to the law of reflection for mirrors. Beam deviations can be easily calculated by use of ray tracing methods. The deviated beam will blur the image quality, enlarge the focus point, degrade the accuracy of the wave-front, and harm the performance of the optical system. To avoid such errors, surface figure must be manufactured and measured precisely.

"If you can not measure it, you can not improve it." This famous quote, attributed to Sir William Thomson, Lord Kelvin, describes exactly the dilemma faced by manufacturers of precision optical surfaces. It is especially true in the fabrication of mirrors used at extreme grazing incidence angles to focus x-rays in telescope and synchrotron applications. Mirrors used in synchrotron beam lines in the 1970s and '80s suffered from excessive surface roughness, generating excessive amounts of scattered light which reduced focal spot intensity and produced spectral contamination in VUV monochromators. The diamond stylus profilometers available at the time were ill-suited for measurements of surface roughness on full-size optical components. The development of the non-contact interferometric phase-measuring microscope early in the 1980's made surface roughness measurements in the Angstrom range on full-sized mirrors possible, and improvements in surface finishing techniques were quick to follow (Bhushan et al., 1985; Koliopoulos et al., 1980; Wyant et al., 1984)

The traditional measurements of plane and spheres applied test plate as a reference to test the surface with interferometric method in early years, which can reach  $\lambda/10$  or better accuracy. However, this is contact test, and they are difficult to achieve with low-coherence sources prior to the development of the HeNe laser. See the excellent discussion of various historic interferometers in Malacara's book *Optical Shop Testing* (Malacara, D., 1992). Later the development of different kind interferometers and the application of interference laser promise the non contact measurements. Computerized interferometer with the CCD sensor application and rapid fringe analysis enhanced the measurement technology accuracy



quickly. Computer generated holograph (CGH) method (Pruss, 2008) can replace the troublesome preparation of test plates. In recent decades the development of the phase measuring interferometer (PMI) becomes an excellent metrology method to promote the interferometer accuracy and repeatability significantly, which has become a standard procedure for the interferometers. The basic concept of the PMI is that one can calculate the precise phase by acquiring several phase shifted frames of interferograms, each phase is shifted by a certain amount. Measurement method accuracy of the PMI could be  $>1/1000$  fringe comparing to previous fringe analysis accuracy of only  $1/10$  fringe. The advantages of the PMI are the high accuracy can be obtained with low contrast fringes and is independent of intensity variations across pupil. Measurement of  $\lambda/20$  surfaces is no longer a black art practiced by master opticians but is possible by ordinary shop technicians using modern phase measuring interferometers. These PMI instruments are widely used in the workshop and the research laboratory because of their ability to make non-contact measurements over large areas with high precision.

The PMI has successfully resolved two major tasks for precise metrology of routine or research surfaces: a) measuring surface roughness with height resolution of less than 1 Angstrom; b) measuring larger 2D surface profiles with  $\lambda/1000$  repeatability and improving the test accuracy. In a conventional phase-shifting laser interferometer, multiple frames of data are acquired in sequence, so there is enough time for vibration and turbulence to degrade the measurement results. Recently developed technology acquires all phase data simultaneously. This will be very useful for in situ precision testing in the workshop (Zecchino, M., 2008; ESDI, 2011)

However, in the case of using PMI, a reference mirror is always required as the measurement standard, which dominates the final possible measurement accuracy.

The metrology problem is more acute in the fabrication of aspheric optics. In recent decades, high technology developments of computer numerical control (CNC), diamond turning, magnetorheological finishing (MRF), ion polishing, and elastic emission machining (EEM) have removed many of the difficulties involved in the manufacture of aspherics and high accuracy conventional optics. These manufacturing techniques have enabled the fabrication of nano-radian and nanometer accuracy components required in various applications.

Several recently-developed metrology techniques extend the measurement capabilities of conventional interferometers. Liu, et al., developed a sub-aperture technique for measuring on-axis aspheres by combining annular regions measured with a Twyman-Green interferometer (Liu et al., 1988). This technique works by adjusting the distance to the test surface to match the wavefront curvature to the surface radius at each zonal region. Sub-aperture stitching interferometry (SSI™), developed by QED, enables the testing of larger aperture optics with standard Fizeau reference optics, without the need for dedicated large null optics, by automatically combining multiple overlapping sub-aperture measurements to form a full-aperture measurement (Fleig et al., 2003; Murphy et al., 200; 2003). The QED system simultaneously produces an error map of both the larger test surface and the smaller reference surface, correcting for rigid body alignment in each sub-aperture and for reference surface errors. Measurement error repeatability has been demonstrated to be 2 nm rms. Relative angle determinable stitching interferometry (RADSI) was also developed by the Osaka University group to measure steeply curved X-ray mirrors with nanometer accuracy

(Mimura et al., 2005; Yamauchi et al., 2003; Yumoto et al., 2010). In this method, the relative angles between overlapping sub-apertures are determined simultaneously with one interferometer while acquiring the sub-aperture profiles with another. Because stitching analysis eliminates certain systematic errors inherent in large aperture systems, the SSI enhances the accuracy of measurements. However, stitching methods can accumulate small systematic interferometer errors that limit very high accuracy. Hence, the measurement accuracy of a full-size surface needs to be calibrated precisely.

## 2. New demands on nano-accuracy metrology

Nanometer and nanoradian accuracy (for simplicity we use “nano” for nanometer or nanoradian or both) is required in many high technology areas: synchrotron radiation (SR) optics, extreme ultra-violet lithography (EUVL), cameras, telescopes, X-ray free electron lasers (XFEL), laser ignition facilities, freeform optics and so on.

### 2.1 Nano-accuracy requirement of synchrotron optics

Hard X-rays produced by synchrotron radiation (SR) sources are important tools for chemical, elemental, and structural analyses of matter at the nano- and atomic scale and for elucidating the molecular processes involved in biological functions at the cellular level. Scientists anticipate having one-nanometer probe spots for such research. Recently, construction started on ultra-bright SR sources with nano-focusing spots at the National Synchrotron Light Source II (NSLS II), Brookhaven National Laboratory (BNL) in the United States and at the Taiwan Photon Source (TPS) of the NSSRC. The NSLS II will allow researchers to create high-contrast X-ray images of matter at this resolution. To focus the bright, hard X-rays of a SR beam into a 1nm spot, beam lines must incorporate a series of precise optical elements. One of the most promising approaches to do so is the application of total reflection mirrors with their exceptional characteristics of broadband focusing, achromaticity, and high efficiency. Osaka University has focused hard x-rays to a spot size of approximately 7 nm using Kirkpatrick–Baez mirrors (K-B mirror) (Yamauchi et al., 2011). However, the mirror figure employed to focus a SR beam to a nanometer spot while preserving coherence requires nano-radian (nrad) accuracy. According to simple geometric-optics calculations, if the error in surface slope is 100 nrad, the beam will exhibit a lateral displacement of 4 nm at a 20mm focal distance; hence, it will greatly enlarge the 1nm spot.

Synchrotron radiation optics, operating at extreme grazing incidence angles on the order of a few milliradians, utilize surfaces consisting of planes, spheres, and aspheres including cylinders, toroids, paraboloids, and ellipsoids, up to 1.5 meters long. These lengthy cylinder-like aspheric surfaces normally have a long tangential radius of curvature, in the range from one hundred meters to several kilometers, with a sagittal radius of curvature that can be as short as a few centimeters. These surfaces are extremely difficult to measure with a traditional null interferometer, even with special computer-generated hologram (CGH) null lenses, or with an SSI. The new beam lines at the NSLS II will require many optical surfaces with 100 nrad slope error.

Fig. 1 is an example of SR optics, a K-B mirror set used for focusing X-ray beam to a nanometer spot. Two nano-accuracy elliptical cylinder mirrors are used to focus a beam to a single point in horizontal and vertical direction individually.

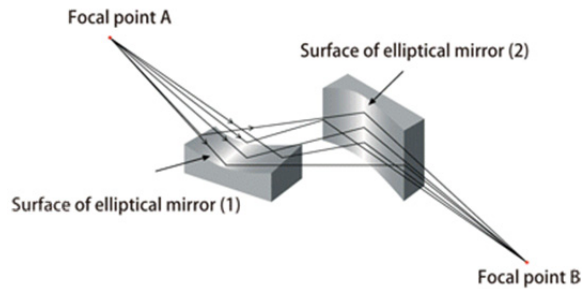


Fig. 1. K-B elliptical mirrors for nanometer spot focusing

Control of surface slope error is especially important in the fabrication of far off-axis aspheres used in grazing incidence x-ray applications. Incoherent x-rays act much like bullets in reflecting off of aspheric surfaces. Slight deviations in surface slope, on the order of fractions of a micro-radian, are sufficient to reduce the quality of the focal spot in long synchrotron beam lines. The stringent requirements for these precision optical components have stimulated research into new measurement methods for spheres and aspheres, and new sophisticated optical measurement instruments are rapidly evolving. A specially designed nano-accuracy profiler for synchrotron optics is urgently needed. It should be also able to measure mid-spatial frequencies and be able to test strongly curved surfaces.

Since the early 1980s, when the first and second-generation synchrotron light sources came on-line, the requirement for RMS slope error tolerance of SR optics has dropped from  $5 \mu\text{rad}$  to  $100 \text{ nrad}$  as the desired focal spot size at the beam line end-station has gone from  $10 \mu\text{m}$  to  $1 \text{ nm}$  (Table 1). These continual machine improvements have driven the development of the metrology instrumentation from the micro-accuracy level to the nano-accuracy level.

Years	Slope error requirement	Spot size
1980s	$5 \mu\text{rad}$	$10 \mu\text{m}$
1990s	$1 \mu\text{rad}$	$1 \mu\text{m}$
2011 for NSLS II	$100 \text{ nrad}$	$1 \text{ nm}$

Table 1. The requirements of slope error and spot size

## 2.2 X-ray free electron lasers (XFEL) and EUVL Lithography

X-ray free electron lasers (XFEL) and EUV lithography require a variety precise of optics with nano-radian accuracy as well. For example, “elliptical mirror with  $0.21 \mu\text{rad}$  rms accuracy is required for XFEL oscillator” (Kim, K., 2009) and “ $0.1 \mu\text{rad}$  accuracy mirror is required for FEL focusing” (Assoufid, L. et al.; Minura et al., 2008). In the case of extreme ultra violet lithography (EUVL) “Neither the figure errors nor the roughness of mirrors for the imaging system must exceed a few angstroms” (Soufli, R., 2011).

## 2.3 Astronomy telescope science requires nano-accuracy optics in large scale

In order to improve the resolution of astronomical telescopes, the desired tendency is to enlarge the aperture of the telescope and achieve the diffraction limited image by use of

perfect optics. This leads to the application of optics with nano-accuracy. This is the new challenge to manufacturing and metrology. Following quoted contents is the clear examples for describing these demands.

“The NASA Science Missions Directorate seeks technology for nano-accuracy metrology on NASA SBIR AND STTR 2011 Program Solicitations. Following are some requirements for telescopes: “Metrology instruments should have 10 nm or better surface height resolution and span at least 3 orders of magnitude in lateral spatial frequency for Optics Manufacturing and Metrology of Telescope Optical Surfaces” and “In situ metrology systems that can measure optics and provide feedback to figuring/polishing instruments without removing the part from the spindle”. “NASA is preparing potential future space telescopes, which have very specific mirror technology needs. UV/optical telescopes require 1 to 3 meter class mirrors with < 5 nm rms surface figures. IR telescopes (such as SAFIR/CALISTO) require 2 to 3 to 8 meter class mirrors with cryo-deformations < 100 nm rms. X-ray telescopes (such as GenX) require 1 to 2 meter long grazing incidence segments with angular resolution < 5 arc-sec down to 0.1 arc-sec and surface micro-roughness < 0.5-nm rms.” (NASA NASA SBIR and STTR 2011 Program Solicitations)

#### **2.4 Precision measurement requirements for large radius of curvature**

In order to ensure similar performance for each laser line of National Ignition Facility (NIF), it is necessary to keep each surface radius of curvature identical. The interferometric method is generally used to measure the radius of curvature of a spherical optical surface. However, it is difficult to perform on surfaces with large radii exceeding a few meters with large aperture, because it requires the test part to be moved over a distance equal to its radius. It means the distance between the surface under test and the reference is very long so it increases the measurement uncertainty. The NIF developed a method, based on dual-focus zone plates to solve this problem (Wang, Q. et al., 2008). In contrast to the NIF, Chengdu Fine Optical Engineering Research Center in China had made the measurement of long radius of curvature on large aperture lens with pencil beam scanning LTP-MF (Ye & Yang, 2011) with very promised accuracy. The nano-accuracy measurement can increase the test accuracy of radii of curvature significantly.

### **3. Early developments in nano-accuracy metrology**

#### **3.1 Scanning profilers**

Except for the PMI and SSI, various specialized optical metrology techniques have been developed over the years to measure this class of optics, based upon scanning profilers: the Random Devices slope scanner (DeCew et al., 1986), the Zeiss M400 CMM in Germany (Becker et al., 1987), various instruments at the National Physical Laboratory in the UK (Ennos et al., 1982; Stedman et al., 1979), and the fringe scanner developed by Hughes Aircraft for the measurement of the AXAF (Chandra) x-ray telescope optics (Sarnik & Glenn, 1987) are examples of scanning profilers. These instruments employ various kinds of metrology methods, contact stylus and non-contact optical, developed to suit the particular metrology problem at hand. The Stedman-Stanley profiler (Stedman et al., 1979) and the Heynacher profilers (Heynacher & Reinhardt, 1979) were stylus instruments that made contact with the surface under test. These instruments made it possible to assess the height of steep aspheric surfaces at the nanometer level in two dimensions. Nevertheless, they are contact measurements, and

during the final testing of optics with delicate coatings applied, there is always the possibility of leaving a very slight mark on the reflecting surface. Non-contact profiling methods have superseded contact stylus methods in the production of most grazing incidence optics today. Profiling instruments based upon the pencil-beam interferometer (PBI) are now widely employed in the manufacture and testing of SR and x-ray telescope optics (von Bieren, 1983a; 1983b; Takacs et al., 1988; Takacs et al., 1989). The PBI has been shown to be quite versatile and ideally suited for the testing of large cylindrical aspheres with nano-accuracy.

### 3.2 Pencil beam scanning surface profiler development in past 25 years

Soon after the NSLS x-ray ring was commissioned in the early 1980's, it was realized that the quality of the focused beam was being compromised by the less-than-perfect figure of the mirrors. The normal-incidence surface figure measurement techniques available to manufacturers at the time were inadequate for characterizing the slope errors in the grazing incidence optics, slope errors that produced blurring of the focus by several times the expected source-size-limited spot size. Conventional interferometry was practically useless in attempting to extract useful surface error information from extremely foreshortened apertures seen at grazing incidence angles on the paraboloids, toroids and ellipsoids in beam line instruments, making it difficult for manufacturers to produce good quality optics. Null lenses did not (and still do not) exist to allow the use of Fizeau interferometers to easily measure simple cylinders for use at 3 mrad grazing incidence angles. With these limitations in mind, we sought to develop a versatile measuring instrument that did not require the use of null lenses and that could handle a wide variety of grazing incidence optics: long, flat mirrors; long-radius spheres; cylinders; toroids; ellipsoids and elliptical cylinders; and bent optics with radii of curvature in the range of kilometers down to tens of meters.

#### 3.2.1 Original pencil beam profiler

Fortunately, around this time period in the early 1980's, the pencil beam interferometer was developed by von Bieren at Rocketdyne (von Bieren, 1983a; 1983b). His system was designed to measure the profile of conical axicon mirrors with a long radius in the tangential direction and an extremely small sagittal radius that vanished at the tip of the axicon. A schematic diagram of his optical system is shown in Fig. 2. We recognized that this slope-measuring profilometer would be ideally-suited for solving our grazing incidence mirror metrology problem.

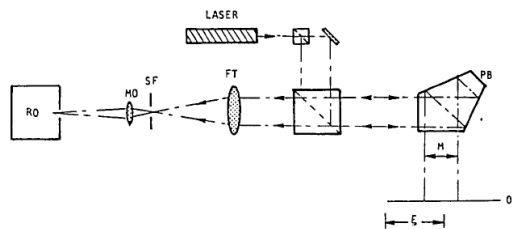


Fig. 2. Schematic of original pencil beam interferometer from von Bieren (vonBieren, 1982). The minimum separation distance between probe beams,  $M$ , is set by the size of the cube beam-splitter and mirror. The beam is scanned by moving the penta prism. A dual lens system projects the interference spot onto the detector.

### 3.2.2 The Long Trace Profiler LTP I and LTP II

The Long Trace Profiler (LTP), based on the principle of the pencil-beam interferometer, was developed by Takacs and Qian and collaborators for the metrology of second-generation synchrotron radiation optics (Takacs et al., 1987; Takacs & Qian, 1989). Its operating principle is similar to that underlying an autocollimator, but with a laser pencil beam employed to scan the mirror being tested. The pencil beam is usually the direct output from a collimated laser diode or fiber-coupled laser. A schematic of the LTP optical system is shown in Fig. 3. The first beam-splitter produces a colinear pair of beams separated by a variable distance set by the adjustable prism. The separation distance can be adjusted from full overlap,  $M=0$ , to any desired value, while maintaining zero optical path difference (ZOPD) between the beams. The ZOPD system design is a vital improvement for a successful PBI profiler, because it eliminates the interference fringe movement due to the laser frequency shift (Qian & Takacs, 2004). This increases measurement accuracy significantly even when using an unstabilized laser. In contrast, the original von Bieren design had a very large OPD between the two beams. The ZOPD system allows the beams to be adjusted with a separation distance equal to the nominal 1mm beam diameter for maximum spatial frequency range information. The beam pair passes through the polarizing beam-splitter, PBS, and is split into a reference beam (REF) that is directed horizontally to a stationary mirror and a test beam that is directed down to the test surface. The return beams are focused onto a linear array detector by a Fourier transform lens, where each beam pair produces an interference pattern adjusted to have a minimum in the center. The position of the minimum on the detector,  $y$ , is proportional to the local slope of the surface between the two components of the beam pair:

$$y = F * \tan 2\alpha \quad (1)$$

where  $\alpha$  is the local surface slope and  $F$  is the focal length of the lens. By scanning the beam across the mirror surface, the slope profile is measured, from which the height can be derived by integration. In the case of synchrotron optics for 2nd and 3rd generation machines, the slope error profile is a more useful measure of the surface quality for the end user, but the height profile is necessary for the manufacturer to use in correcting the surface.

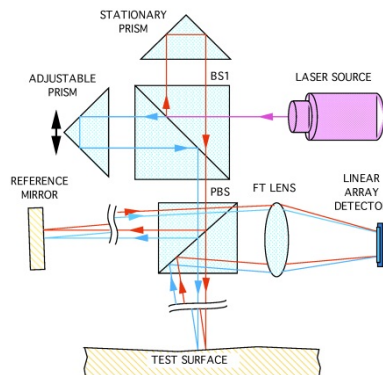


Fig. 3. Schematic of the LTP-II optical system. The PBS generates a reference beam from a plane mirror fixed to the optical bench surface. Pitch error in the movement of the optical head on the air bearing is corrected by adding the signals from the test and reference arms.

Some recent versions of the LTP use a single probe beam instead of the dual beam, eliminating the need for the initial beam splitting optics. The probe beam is focused to a single spot with no internal interference structure. The peak or centroid of the intensity distribution on the sensor determines the angle of the surface seen by the beam footprint. In this respect, the LTP operates very much like an autocollimator. Various algorithms are used to extract the centroid location of the spot with high precision.

Over the years, many improvements have been made to the original LTP-I design by many collaborators involved in synchrotron and x-ray telescope optics metrology. The original LTP system utilized an external electronic autocollimator to measure the pitch angle error of the optical head as it moved along the air bearing. The autocollimator was replaced by the internally-generated reference beam, shown in Fig. 3, by Irick, et al. (Irick et al., 1992) This allowed for simple correction of the measured profile for mechanical pitch errors by adding the two signals together. The relative intensity between the test and reference beams could now be adjusted easily by the use of polarizers and wave plates added to the optical system (not shown). The commercial version of the instrument, the LTP-II, produced by Continental Optical Corp. for several years, incorporated the internal reference beam and used a dual-array linear photodiode sensor as the detector. The dual array detector allowed the reference beam to be aligned nearly along the optical axis of the system instead of at an extreme angle, which places the spot at the end of the sensor. Having the reference beam centered in the lens aperture minimizes the introduction of phase shifts by glass inhomogeneities that translate into beam spot location variations with lateral movement of the beam across the aperture during a long scan. Bresloff added a Dove prism into the reference beam to change the phase of the laser pointing direction drift to be the same as the mechanical pitch angle error, allowing for correction of both error signals simultaneously by addition of the two signals (Takacs & Bresloff, 1986; Takacs et al., 1999).

#### LTP II OPTICS BOARD

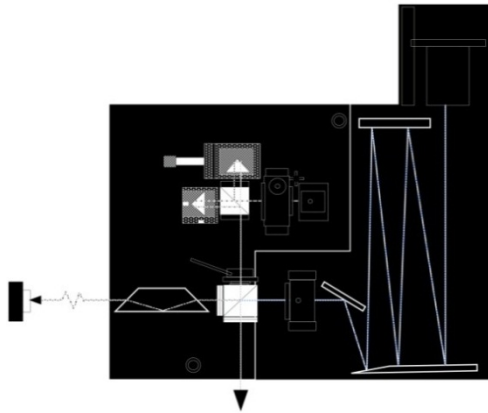


Fig. 4. The LTP II optics board layout showing the 10 mrad surface slope acceptance angle optics in place. A 4-mirror arrangement folded the beams from the 1.25 m focal length lens with 7 reflections onto the detector mounted in back of the plane of the figure. The Dove prism mounted on the optics board in the REF arm inverted the phase of the pitch error signal to allow for simultaneous correction of pitch and laser drift

A sketch of the LTP-II optical head is shown in Fig. 4. This optical head used a Fourier transform lens with a focal length of 1.25 meters, which necessitated a folding mirror system to keep the size of the system within reason. The system was designed to measure surfaces with a total angular range of 10 milliradians. While this may seem like a small acceptance angle, it was sufficient to handle 99% of the long-radius optics used in NSLS beam lines. The LTP excels at measuring large flat and long-radius surfaces up to 1 meter in length. Other versions of the LTP-II can handle mirrors up to 1.5 meters long.

Improvements to the LTP-II and its successors, the LTP-IV and -V, manufactured by Ocean Optics, eventually enabled reliable surface slope error measurements down to the 0.5  $\mu\text{rad}$  RMS level. However, recent advancements in synchrotron machine technology have resulted in the need for mirrors with slope errors in the 100 nrad range in order to allow for nanometer focal spot sizes. This quest for nano-accuracy in metrology has led to the development of specially-engineered machines that must be used in thermally-stable special environments in order to achieve this level of accuracy. The NOM machine developed at BESSY II in Germany by Lammert and collaborators (Lammert et al., 2006; Siewert et al., 2004) is the prime example of this next-generation profiler, which we will discuss in the next section.

### 3.2.3 Penta-prism LTP

The LTP II uses a scanning optical head (OH) in the tilted reference mode. Qian et al. developed the scanning penta-prism mode LTP (PPLTP) in 1995 at Sincrotrone Trieste, Italy (Qian, 1995), an evolution that extends applications and improves the accuracy of tests for plane- and near plane-mirrors because a tilted reference beam is unnecessary. The main characteristic of the penta-prism is: the angle between incident and output beams of the penta-prism will always be equal to  $2\alpha$  ( $\alpha$  is angle between two reflection surfaces of the penta-prism, nominal angle is  $45^\circ$ ), even if the penta-prism is tilted. So for a scanning penta-prism, the slide pitch error will not influence output beam direction. This measurement method has been successfully adopted by many SR metrology laboratories. In addition to accuracy improvement, the penta-prism LTP has enabled the testing of in situ heat load distortion of mirrors in vacuum chambers and the testing of small diameter aspheric surfaces of astronomy telescopes by use of a derivative of the penta-prism LTP: the in-situ LTP and vertical scan LTP (VSLTP). These will be described later. Also, the NOM employs a scanning penta-prism mode to enhance accuracy to the nano-radian level in small test angles.

### 3.2.4 The multiple functions LTP (LTP-MF)

Qian has developed a Multiple Functions LTP (LTP-MF) that incorporates two compact LTP optical heads into various configurations that allow for self-correction of scan-induced errors with subsequent improvements in measurement accuracy (Qian et al., 2005; Qian & Takacs, 2007). Some important facts of the LTP-MF, related to the approach to nano-accuracy, are briefly described here. The LTP-MF can operate in the mode of scanning optical head with non-tilted reference by use of a second optical head to ensure higher accuracy (Fig. 5). If a high quality air-bearing system is used and  $0.01^\circ\text{C}$  temperature stability can be maintained, the LTP-MF can achieve  $0.1\mu\text{rad}$  rms accuracy in testing plane mirror surfaces. The LTP-MF can also operate in the penta-prism scanning mode for testing nearly plane mirrors with high accuracy.



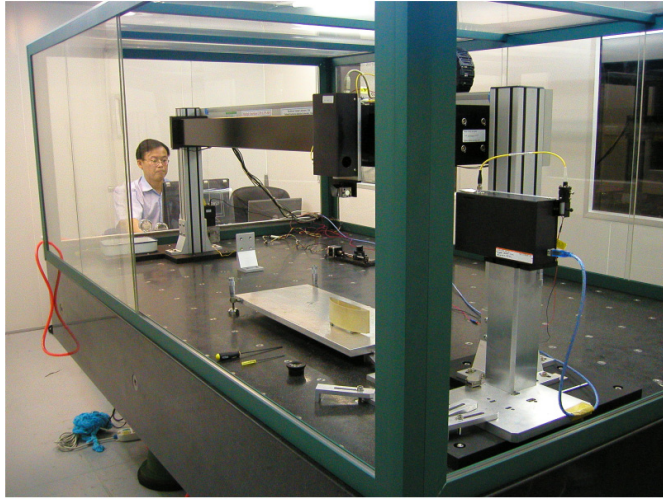


Fig. 5. The LTP-MF collaborated between BNL (USA) and NSRL (China). The first optical head (OH) acts as scanning OH with non-tilted reference by use of the second OH. The second OH can also act as scanning penta-prism LTP

The schematic of the LTP-MF optical head is shown in Fig. 6. An unstabilized diode laser (DL) with a 633nm wavelength with power of 1-5mW is used as the light source. The optical fiber is used as a controllable beam transport tool. There are several advantages to using an optical fiber: a) to achieve a compact and convenient package; b) to isolate the thermal source from the optical system, which is helpful for achieving nano-accuracy; c) to minimize laser beam pointing error; d) to change the light source wavelength with ease and to replace the source without distorting the beam direction; e) to insert optical fiber devices in order to perform various new functions, for example, adding a fiber attenuator for controlling intensity. The beam is collimated to 1 mm diameter spot by a fiber collimator (FC). Then it passes through a monolithic wave-front splitting beam splitter (WSBS) and becomes two half-beam spots with a  $\pi$  phase-shift so that when it is focused on the CCD by lens (FT), it will produce an interference fringe with the shape of valley-at-center for low noise fitting. The LTP-MF uses an equal optical path WSBS for the purposes of nano-accuracy and compactness. Otherwise, the large frequency shifts of diode laser will produce the fringe position shift, which will degrade the nano-accuracy. A pair of microscope cover plates are constructed as an adjustable phase shift WSBS, which is the easiest and lowest cost solution (Qian & Takacs, 2004; 2003)

The 20 x 20 mm PBS splits the beam into sample and reference arms, then they are reflected back from the mirror under test and reference mirror to the FT lens which has a focal length of 400 mm and a 28 mm aperture. The return beams are then focused onto a linear array CCD of 14  $\mu\text{m}$  pixel size. Two quarter wave-plates (QWP) are for isolating unwanted reflected beams, the half wave plate (HWP) is for adjusting the intensity ratio between the sample beam and reference beam. However, in the case of using the scanning OH with non-

tilted reference, the HWP is unnecessary. Polarizer (P) is for adjusting the beam intensity, but now it is easily replaced by a fiber attenuator. Two folding mirrors are for reducing the overall mechanical length, but for new nano-accuracy profile they should be removed in order to reduce systematic error. The FT lens is designed to keep aberrations below  $1 \mu\text{rad}$  for two scanning modes conditions. The PBS needs to be of extremely high quality, including each surface, angles, and material uniformity to insure LTP precision. Good alignment of the optical system will ensure measurement accuracy when the angular test range is large.

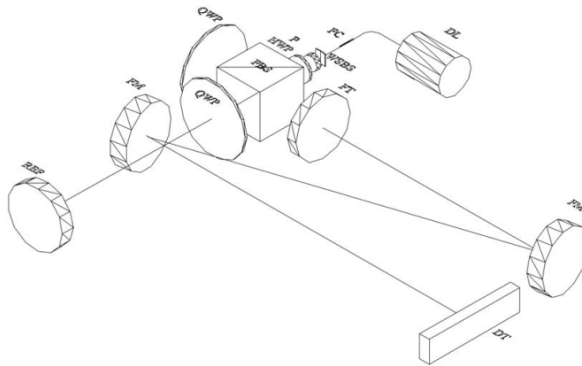


Fig. 6. The optical system scheme of the multiple functions LTP (LTP-MF)

It is conceptually easy to enlarge the LTP measurement range by reducing the focal length of the FT lens, by increasing the CCD size, by using a high resolution CCD, or by enlarging the PBS and FT aperture size. However, if nano-accuracy is required over the entire angular test range, then the improvement task becomes extremely difficult.

### 3.2.5 The vertical scanning LTP (VSLTP)

The mirrors used for X-ray telescopes are all based upon the use of grazing incidence optics in various configurations. Wolter Type I systems use a combination of paraboloidal and hyperboloidal surfaces; a Wolter-Schwarzchild Type I system consists of a small figure modification of a Wolter I system; the foil cone is an approximation of a Wolter I system, and the Kirkpatrick-Baez (K-B) system consists of two sets of orthogonal spheres or parabolic cylinders, produced by bending thin plates. The surface figures are generally conical in shape with very small sag deviations, on the order of microns, from the best-fit conical surface. These optics are ideally suited for measurements with the LTP. The ideal configuration for measuring x-ray telescope optics, however, is to measure the object while it is oriented with its symmetry axis in the vertical direction. This minimizes the effects of gravity-induced distortion on the surface figure, especially on thin shells or foil surfaces. A vertical scanning LTP (VSLTP), a modification of the scanning penta-prism LTP, was developed for testing X-ray telescope mirrors in the vertical orientation and is shown in operation in Fig. 7 (Li et al., 1996,1997). The benefit of the VSLTP is that a small penta-prism can be scanned inside a small diameter x-ray telescope mirror.

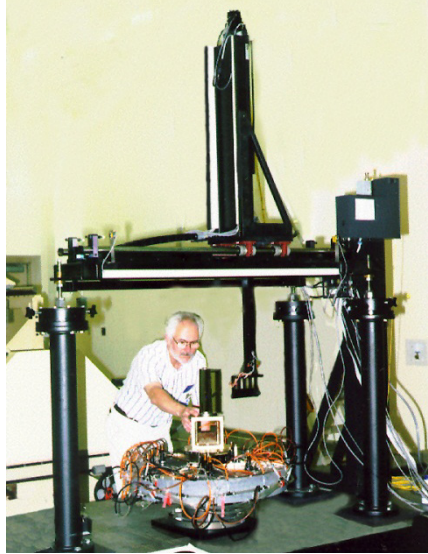


Fig. 7. Vertical Scan LTP (VSLTP) at Marshall Space Flight Center set up for measuring x-ray telescope mirrors and mandrels in the vertical orientation

In 25 years of development, there have been a number of other applications of the LTP: in-situ heat load test, measurement at machine shop, calibration of the profiler, thermal shift treatment, 2D detector development and so on. Some will be described in the following sections.

#### **4. Current development and trends in nano-accuracy surface figure metrology**

Precise metrology is the basis for enabling fabrication of precision optics. The rule-of-thumb requires that the metrology accuracy should be at least 3-fold ( $< 50$  nrad) better than the specification of the optics. The great challenge for metrology and manufacturing is that nanometer and nano-radian accuracy is needed for spherical or aspheric mirrors with large surface slope angles.

##### **4.1 Pencil beam scanning method for nano-accuracy surface figure measurements**

Progress in nano-accuracy metrology is dependent upon new scientific demands, advanced technology developments, and innovation in metrology methods and metrology instruments.

Though traditional null phase-shift interferometers can reach excellent repeatability with high accuracy, the required reference surface is still an obstacle, which restricts its metrology accuracy. Making a null lens is time-consuming work and is very expensive, especially for one used to test large-aperture optical surfaces. Though the three flats absolute calibration method can reach an accuracy  $\lambda/100$  or better theoretically (Schulz et al., 2008), it is very difficult to achieve for large surfaces. In addition, the reference surfaces have to be routinely

but inconveniently calibrated. As described above, the stitching methods accumulate small systematic interferometer errors that limit their high accuracy. Hence, the accuracy of measurements of a full-size surface with sub-aperture stitching method needs to be calibrated precisely

Pencil beam scanning methods have become a promising metrology method to make nano-accuracy surface profilers, because of their many advantages: non-contact test, absolute measurement without need of using a large reference, high accuracy, possibility of measuring large dimension optics and aspheric optics with moderate cost, and no need for working distance adjustment due to the use of a collimated beam. These advantages keep this method as one of most important solution for the optical metrology in the future even though it has a disadvantage of being only a one-dimension measurement with lower test speeds. The sub-aperture interferometer stitching method is beginning to be used with LTP and SR-optics measurements (Assoufid et al., 2004; Polack et al., 2010) with linear scanning. It will extend to large optics the ability to perform 2-D and mid-spatial frequency testing. However, improvements are needed in LTP stitching accuracy, along with research on precise calibration methods.

#### 4.2 Various nano-accuracy profilers with scanning pencil beam method have been developed

The Nano-Optical-Measuring Machine (NOM, Fig. 8) (Lammert et al., 2006; Siewert et al., 2004) is the most accurate instrument so far for evaluating SR and other large optics. The NOM incorporates a special commercial autocollimator and an LTP optical head with a scanning penta-prism system to measure long-radius optics. It applies a small aperture of about 2 mm near the mirror under test in order to increase the spatial frequency range. The demonstrated uncertainty of the NOM in the measurements was low: for a plane mirror it was  $0.05 \mu\text{rad rms}$ , and, for curved mirror, it was  $0.2 \mu\text{rad rms}$ . Both instruments operate in a scanning penta-prism mode without the need to use a reference beam to correct for slide pitch error. A similar and improved Diamond-NOM has also been developed (Alcock et al., 2010).].

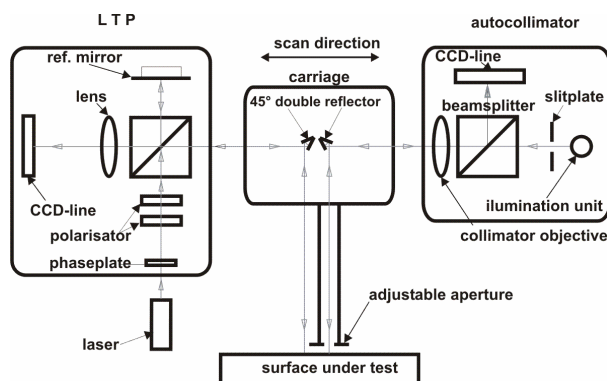


Fig. 8. The NOM, Courtesy of F. Siewert, T. Noll, T. Schlegel, T. Zeschke, H. Lammert, The nanometer optical component measuring machine: a new sub-nm topography measuring device for x-ray optics at BESSY, AIP CONFERENCE PROCEEDINGS Vol.705, 847-850 (2004)

A new Traceable Multiple Sensor (TMS) system was developed by the Physikalisch-Technische Bundesanstalt (PTB) for measurement with nanometer accuracy (Fig. 9) (Schulz et al., 2010; Wiegmann et al., 2010; PTB Working Group 8.42, 2011). It encompasses coupled multiple distance-sensors that are scanned along the surface under test. By using a small sensor head, a high lateral resolution is achieved. In addition to the multiple distance sensors, the TMS utilizes an autocollimator measuring the tilt of the sensor's head, thereby eliminating systematic errors in the distance sensors. The TMS can reach nanometer or better accuracy with high lateral resolution. Both the NOM and TMS are suitable for plano and near-plano mirror measurements with nano-radian accuracy. During a scan on a plane mirror, the reflected beams are always parallel and remain steady except for very tiny angle variations produced by the slope error on the surface. In this fixed beam direction condition when testing plane mirrors, if the temperature is very stable and the scan system has very small pitch, yaw and roll errors, the beam will not have any error that impacts the test accuracy except for the noise. So it is possible to achieve nano-accuracy on plane surfaces regardless of what instrument is used.

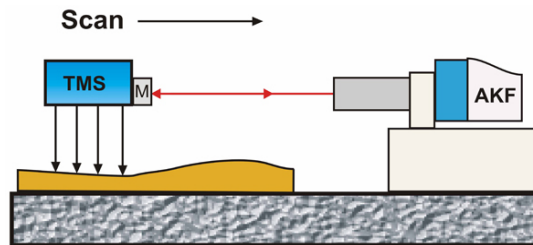


Fig. 9. Traceable multiple sensor (TMS) system, Courtesy of Schulz, G. Ehret, M. Stavridis, C. Elster, Concept, design and capability analysis of the new Deflectometric Flatness Reference at PTB, Nuclear Instruments and Methods in Physics Research A 616 (2010) 134–139

#### 4.3 Difficulties to approach nano-accuracy metrology for large slope surfaces

Testing of strongly curved surfaces presents significant difficulties to pencil beam scanning profilers, as it does to most optical measuring techniques. In this case, measurement nano-accuracy is hard to reach in large part due to the impact of insufficient optical system quality of the profiler.

Let us analyze the beam position variation in the optical system during the measurements. We use as an example a spherical mirror under test (MUT) scanned by the scanning optical head of the LTP II (Qian, S. N. & Qian, K., 2010; Qian, 2011). The sample beam (solid line, Fig. 10 a) measures the slope of a MUT, and the reference beam (dashed line) measures the air bearing pitch error. In order to avoid the overlapping of sample and reference beams on the CCD, the LTP II reference beam is tilted to move the spot to one end of the detector. During the scan both sample and reference beams have lateral motions (BLM) over the optical components inside the optical head, shown as solid and dashed shadow areas, which will pick up large local phase shift errors that show up as surface slope error. These errors are produced by surface figuring error, inhomogeneity of

optical components, system aberration and system alignment errors. However, sample beam lateral motion is an unavoidable condition in testing the slope of curved mirrors, but we can effectively reduce the sample BLM magnitude by adopting a novel system scheme.

In contrast the penta-prism scan mode (by use of the LTP optical head or autocollimator) has much larger BLM than scanning optical head (Fig.10 b), so it is not recommended for measuring larger slope optics.

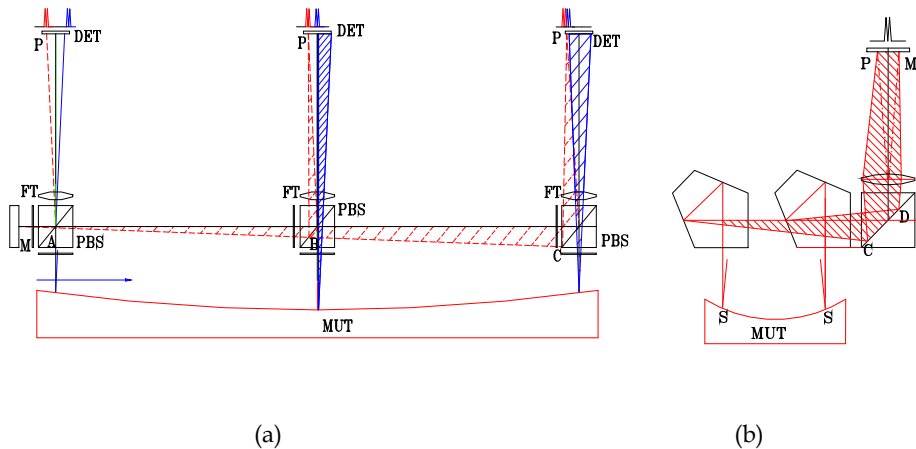


Fig. 10. a) Sample and reference beams' lateral motion in scan optical head mode; b) Beams' lateral motion on a scan in the penta-prism mode

How large is the slope error produced by the BLM? The following measurement compares a tilted reference and a non-tilted reference. Two scans were done only with the reference arm of the LTP III over a 900 mm length scan in sequence. The first scan is with a reference beam angle of 1.5 mrad, which results in a 3 mm lateral motion across the PBS (Fig. 11 a); the second scan is with a non-tilted reference beam (Fig. 11 b). The difference (Fig. 11 c) between both scans is  $\pm 5 \mu\text{rad}$  (P-V) which is a serious slope error for a nano-radian surface profiler. If the tilt angle increases due to strongly curved mirror, the systematic error of the profiler will be more severe. Also, if the reference beam spot is displaced away from system center even in the case of using a second linear CCD or 2D CCD, it will still produce considerable BLM in vertical direction in 2-3 mrad angle level. This method should also not be suitable for the nano-accuracy system. Results of several tests indicate that the magnitude of the slope error caused by BLM could be larger than  $1 \mu\text{rad}$  rms. If the BLM is larger due to the larger test angle, the error will increase quickly. The problem is that the real error is so larger even though we use available highest quality optics, it is still not enough to reach  $0.1 \mu\text{rad}$  accuracy for strongly curved surface test. The following simulation analysis illustrates this effect.

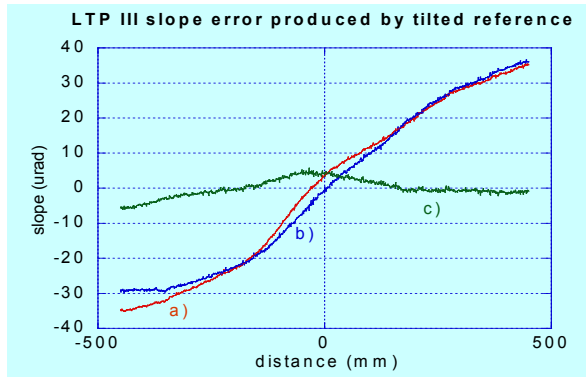


Fig. 11. LTP III slope error produced by tilted reference: a) reference beam spot is tilted in 1.5 mrad; b) non-tilted reference beam at CCD center; and c) slope error caused by tilted reference.

#### 4.4 Error simulation analysis of surface figure errors and inhomogeneity of profiler optics based on wave-front distortion

A sinusoidal wave-front is used to simulate the surface figure error and slope error. A sinusoidal wave-front error of  $\pm 1\text{nm}$  (P-V) in 20 mm (Fig. 13 a) will produce  $\pm 0.314 \mu\text{rad}$  slope error (Fig. 12).

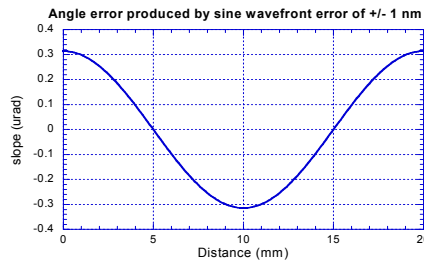


Fig. 12. Angle error produced by sine wave-front error of  $\pm 1 \text{ nm}$

##### 4.4.1 Slope error produced by surface figure error

If a refractive surface figure error is a  $\pm 1 \text{ nm}$  sine wave ( $\lambda/317$ ), it will produce a sinusoidal wave-front error of  $\pm 0.5\text{nm}$  and a slope error of  $0.157 \mu\text{rad}$  for a material with an index of 1.5 (Fig. 13 b), which is a value large enough to destroy the nano-accuracy. If there are multiple surfaces in the optical system, the error will be larger than  $0.1 \mu\text{rad}$ .

If a reflective surface figure error has a sinusoidal error of  $\pm 1 \text{ nm}$ , it will produce a reflective sinusoidal wave-front error of  $\pm 2\text{nm}$  and a slope error of  $0.628 \mu\text{rad}$  (Fig. 13 c), which is a much larger impact than a refractive surface.

The best surface figure quality typically available in customer optics is about  $\lambda/100$  with very high cost. This means that even if we use the highest quality available optics it will probably not be easy to reach  $0.1 \mu\text{rad}$  accuracy.

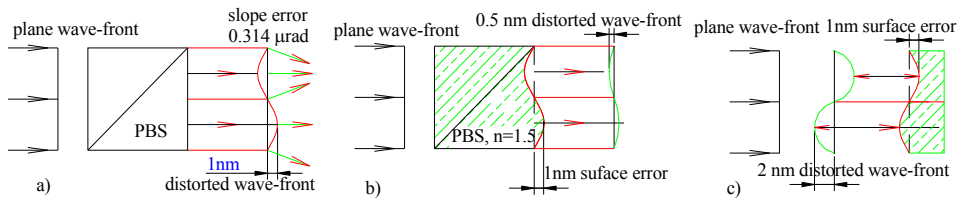


Fig. 13. a) 1nm distorted wave-front and its slope error;  
 b) 1nm refractive surface error and its distorted wave-front;  
 c) 1 nm reflective surface error and its distorted wave-front

#### 4.4.2 The slope error produced by inhomogeneity

The optical path difference (OPD) produced by glass inhomogeneity is

$$OPD_h = \delta n * \text{path length} \quad (2)$$

where  $\delta n$  is the refractive index error. Commercial optics typically use Grade H2 of SCHOTT glass with  $\delta n=5e-6$  (SCHOTT, 2004). Assuming the refractive index variation has a sinusoidal distribution in 20mm length and 2mm thickness layer, it will produce a  $1.57\mu\text{rad}$  slope error. So grade H2 is not suitable for the nano-radian accuracy application. Grade H5 glass or Grade 0AA of Corning HPFS® glass (Corning, 2008) has an index variation of  $\delta n=5e-7$ , which will produce  $0.157\mu\text{rad}$  slope error. Actual error needs to verify with practical test.

Most non-contact optical profilers that use optical systems for measuring angle variation will confront with these difficulties if there is BLM. Examples of such systems are the optical head of the LTP, the autocollimator of the NOM and optical system of the PTB's profilers.

As a result of the above simulation analysis, it is necessary to apply the best quality glass Grade H5 and  $\lambda/100$  figured surfaces to achieve a nano-accuracy profiler. An actual error estimation using the highest quality optics is underway.

#### 4.4.3 Surface figure error in the mid-spatial frequency range and its metrology

Optical surface errors are general divided into three categories: a) low-spatial frequency (LSF): as surface figure error; b) mid-spatial frequency (MSF): as ripple, and c) high frequency: as surface finish roughness (Youngworth & Stone, 2000; Youngworth et al., 2008). The low-spatial frequency surface error is defined over the spatial period range from 5-10 mm to the entire surface dimension, while the mid-spatial-frequency (MSF) surface error is roughly defined in the range between 0.1 to 5-10 mm spatial periods. The impacts to the optical system quality of both LSF and MSF can be analyzed with ray-based model. MSF could also be considered as surface figure.

MSF will have a more severe impact on surface slope error than LSF error. The following sinusoidal wave-front simulations describe this effect. Fig. 14 shows simulated sinusoidal wave-fronts and their slopes (derivative of the wave-front) in a 20 mm surface length. The wave-fronts have identical amplitudes of  $\pm 1\text{nm}$ , equal to  $\lambda/317$ , but with different



frequencies. It is obvious that higher frequency waves have a much larger slope error than low frequency ones.

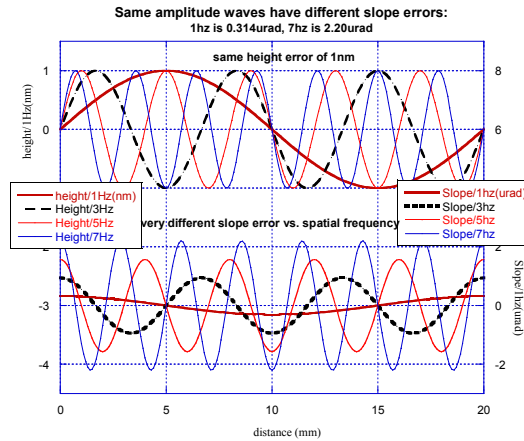


Fig. 14. Higher frequency wave-front has much larger slope error even though the amplitude of the wave-front is the same

Traditional surface figure specification is generally defined as height deviation from a nominal surface figure in number of wavelengths. Traditional optical manufacturing uses long pitch or large lap polishing, which will produce relatively small MSF error. It means fewer ripples on surface. Due to new deterministic polishing technology and metrology accuracy developments, LSF height error can be reduced significantly by removing local material precisely according to precise surface test results. However, local figure correction polishing by use of deterministic polishing may leave small ripples on surface. Though the ripple height may be smaller than  $\lambda/100$ , the short ripple distance (MSF) will have a larger slope than the traditionally polished surface, which could degrade the slope error accuracy. For example, after processing half of the part with fluid jet polishing, the surface flatness P-V is reduced to  $\lambda/100$ , but the MSF ripples (Fig.15 a) could increase the slope error. Fig. 15 b) displays the residual ripples on the surface (Lightmachinery, 2011).

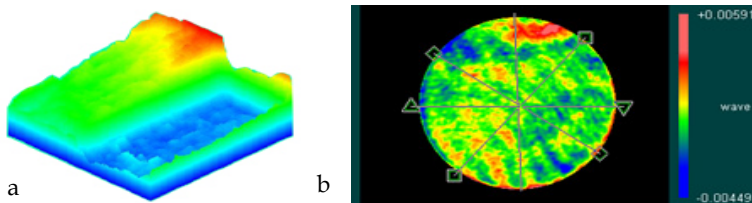


Fig. 15. Fluid jet polishing: a) After processing half of the part with fluid jet polishing. The surface flatness P-V is reduced to  $\lambda/100$ ; b) After fluid jet polishing the 4" silicon mirror has a surface flatness of 1/100 wave (P-V). Courtesy of Light machinery, Fluid Jet Polishing, <http://www.lightmachinery.com/Fluid-Jet-Polishing.html>

Total Surface height error	Slope error without MSF	Slope error with 10Hz MSF
$\lambda/100$	$\pm 1.98 \mu\text{rad}$	$\pm 19.8 \mu\text{rad}$
$\lambda/20$	$\pm 9.9 \mu\text{rad}$	

Table 2. Slope error of 20 mm period sine wave-front with  $\lambda/100$  or  $\lambda/20$  amplitude and different spatial frequency

The above table 2 shows that the slope error of a  $\lambda/20$  surface without MSF could be better than a  $\lambda/100$  surface with MSF. It means that if the MSF can be controlled, the requirement on surface height error can be reduced significantly. Modern sub-aperture and deterministic optical fabrication techniques are more prone to ripple errors. So after deterministic polishing, an additional surface smoothing process may be necessary if there is considerable MSF error. Also, striae are an important material defect that creates MSF and should be considered more carefully. As a result of this analysis, it may be necessary to specify surface error by use of slope error in addition to the use of ambiguous surface height error.

The metrology for the MSF requires higher lateral detection resolution. Point scan surface profilers such as the LTP and NOM apply pencil beam spots of one to several millimeters as a detecting tool, so it is very suitable for the LSF test and will cover partial range of the MSF. Stitching interferometry covers the higher MSF range. However, a new nano-accuracy profiler for large surface test covering the 0.1-2mm MSF range is necessary for high quality optical system, and is currently under development.

#### 4.5 A nano-accuracy solution for a new surface profiler in the large slope testing range: Adopting scanning optical-head combined with non-tilted reference beam

There is increasing demand for nano-accuracy in the testing of strongly curved surfaces, for example, for K-B mirrors used for X-ray synchrotron radiation, for optics for extreme ultra-violet projections lithography and for the new astronomical telescopes.

As described above, there is beam lateral motion (BLM) in the optical system during the beam scan on a curved surface, which will produce significant slope error. If this error cannot be removed or greatly decreased, nano-accuracy will be difficult to reach in the range of large-slope tests.

Good design of the profiler system can reduce BLM significantly and very effectively. The first step is to use as few optical components as possible in order to eliminate unnecessary BLM error. The second step is to design novel optical system of the profiler to minimize BLM. Adopting the mode of scanning optical-head combined with non-tilted reference beam is an effective solution (Qian, 2011) for a nano-accuracy surface profiler (NSP).

The first optical head, fixed on the carriage, is scanned along the air bearing to probe the surface slope, and the second optical head (as used in the LTP-MF or using an autocollimator) is fixed to the granite table for measuring the air-bearing pitch error. In this way its beam can be set without BLM, because there is no problem with sample beam and reference beam spots overlapping. Application of the non-tilted reference will eliminate the BLM completely, so no error will be produced in the reference arm. Other non-tilted reference methods can be used to simplify the system.

The first advantage of applying a scanning optical-head is to create the opportunity to use a short fixed working distance for the sample beam. In this way, the sample beam's BLM can be significantly reduced to  $\pm 1\text{mm}$  (50 mm working distance) in comparison with  $\pm 20\text{mm}$  BLM in the scanning penta-prism mode for a test range of  $\pm 0.01$  rad. This significantly lowers slope systematic error.

The second great advantage of applying a scanning optical head is its very simple calibration. Only one error compensation curve will be necessary to correct for all systematic errors for testing various mirrors. In contrast, it is very hard to compensate systematic errors in the penta-prism scanning mode, in which mirrors with different radius of curvature, different dimensions and different scanning start positions will need different compensation curves. Obviously, for a large test range, using the penta-prism scanning mode effectively precludes reaching nano-radian accuracy due to the BLM.

The third advantage of applying a scanning optical head is that small BLM and fixed working distance minimize the operational aperture of the lens, so it simplifies the aberration-reduction design of the lens. Another necessary approach to reduce systematic error is to improve the quality of the optical components including surface quality, optical material inhomogeneity and roughness.

Scan mode	working distance (mm)	BLM(mm) / at test angle	Test angle (mrad)	Extra optics	Comment
Scan OH+ non-tilted REF (NSP)	Sample: 50 (fixed) Ref: 100-1100	Sample: $\pm 0.5/\pm 5\text{mrad}$ $\pm 1/\pm 10\text{mrad}$ Ref: 0	$\pm 10$	N/A	Larger test angle + high accuracy
Scan OH+ tilted REF (LTP II)	Sample: 50 (fixed) Ref: 100-1100	Sample: $\pm 0.5/\pm 5\text{mrad}$ Ref: $\pm 10/\pm 5\text{mrad}$	$\pm 5$	N/A	Suitable for plane & near plane mirror test
Scan Penta-prism (PPLTP, NOM)	Sample: 300-1300	Sample: $\pm 10/\pm 5\text{mrad}$	$\pm 5$	Penta-prism /mirror	Suitable for plane & near plane mirror test

Table 3. Comparison of three scanning modes

Table 3 is the comparison of three scanning modes

Recently, the PTB described new research in "Scanning deflectometric form measurement avoiding path-dependent angle measurement errors" to reduce the BLM problem (Fig. 16) (Schulz et al., 2010). In the case of the PTB, the first autocollimator (AC1) beam is scanned through a penta-prism to the mirror under test (MUT). But the MUT is no longer stationary now, and it is tilted by a tilting stage in order to direct the reflected sample beam back along the incoming direction. This means that the MUT is tilted according to the slope at each scanning point. So this measurement arm incurs no BLM. The second stationary

autocollimator (AC2) is used to measure the tilt angle/slope of the MUT with a mirror fixed to the tilting stage. In this way, this tilting angle/slope test arm has a fixed short distance, which will reduce the BLM significantly (as in the case of the scanning optical head with non-tilted reference method). This is a good method to reduce measurement error caused by BLM.

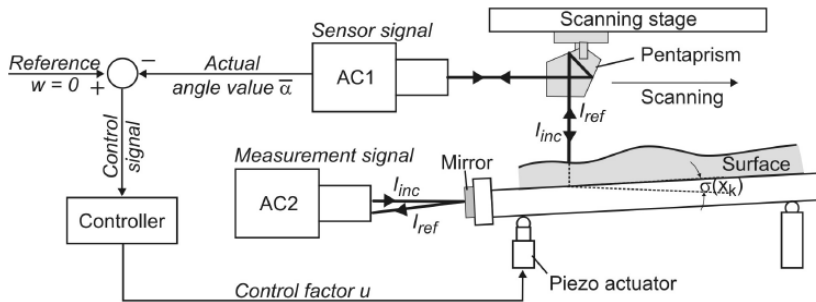


Fig. 16. Principle of operation of the EADS system. AC1: Straightness representation and null instrument, AC2: angle measurement.

Courtesy of M. Schulz et al. Scanning deflectometric form measurement avoiding path-dependent angle measurement errors, JEOS Rapid Publications 5, 10026 (2010)

The main principle and similar strategy in both cases of the scanning optical head with non-tilted reference method and the PTB method are: the probe arms to measure large tilted angles are short and fixed in order to reduce the BLM and are combined with another non-tilted beam arm in order to eliminate the BLM. As a matter of fact, in the PTB case, the angle tilt test is converted to a new arm by use of a tilting stage. However, a precision tilting stage must be used.

#### 4.6 Summary of unresolved problems for nano-accuracy surface measurement

There are still many problems that need to be solved in the development of the NSP.

1. Further improvement in nano-accuracy: a  $0.05 \mu\text{rad}$  accuracy is expected
2. To ensure nano-accuracy ( $0.1 \mu\text{rad}$  rms) in the whole test range of 10 mrad (recent available test range) to 20 mrad (for strongly curved surface, like K-B mirror)
3. Except for stitching method, surface figure metrology of middle-spatial frequency of 0.1mm-2mm is necessary. It means to develop a NSP to span the measurement spatial frequency range from 0.1 mm to 1 meter on one scan.
4. For full 2D testing, now it is necessary to make parallel lines scans and multiple intersectional scans in order to establish the relations for every scanning point. It is better to obtain the slopes in both X and Y directions at every scanning point simultaneously. In this way measurement accuracy and time could be further improved.
5. In situ and cost-effective nano-accuracy calibration system development is necessary

## 5. Calibration of nano-accuracy

As described above, now the required measurement accuracy is nano-radian and nanometer. If the instrument cannot be calibrated precisely, nano-accuracy metrology achievement is meaningless. Plane mirror measurement by use of traditional phase shift interferometer is difficult to reach  $\lambda/50$ - $\lambda/100$  accuracy because of the reference surface accuracy limitations. The absolute flatness test of the plane mirror by use of three-flat test method can reach  $\lambda/100$  or better, but it is very difficult for large mirror calibration (for example for 500mm to 1000mm dimension). Pencil beam profilers can solve this problem, but it is only in 1-D. When the tested slope range increases, the angle related systematic error of the measurement instrument increases significantly. Though the pencil beam profilers have potential power to reach nano-radian and nanometer accuracy or less in principle, they still need to be calibrated. Calibration equipment should have the accuracy of three times better than the instrument to be calibrated. This is a great challenge to optical metrology.

### 5.1 Angle calibration system based on trigonometric function

Sine bar and tangent bar methods can be used to precisely measure small angles.

#### 5.1.1 Sine bar

The sine bar is a simple and effective method in calibrating small angles. Sine bar system contains one solid bar and two fixed-distance cylinders. When one cylinder B is lifted in a height (h), while another cylinder A is kept in contact with the base surface (Fig. 17), the bar rotation angle  $\alpha$  will be:

$$\alpha = \arcsin (h/L) \quad (3)$$

If the roundness and diameter of two cylinders are very precise, and their geometric positions and moving height are accurate, the angle  $\alpha$  can be calculated very accurately. The distance error of two cylinders is a systematic error, which can be eliminated easily by calibration. The sine bar technology is able to reach nano-radian accuracy level.

The sine bar system known as the small angle generator with a length of about 523 mm at National Physical Laboratory (NPL) of Britain is used for autocollimator and small angle calibration (NPL, 2010). Its uncertainty is estimated to be  $\pm 0.03$  second of arc for angles in the range  $\pm 10$  minutes of arc. It is suitable for nano-radian profiler calibration in this angle range. For larger angle calibration the accuracy needs to improve.

The height lifting of a traditional sine bar is made by inserting a precise gauge block under one of the sine bar cylinders, or by use of micrometer. However, the process of cleaning and wringing of gauge blocks and support surface is time consuming and is not very reliable for reaching nano-accuracy. The gauge block accuracy is 50nm. Generally the average wringing film thickness between the gauge blocks is about 10 nm, but some wrings will be over 25 nm. For a non-professional operator the error could be much larger. A 30 nm possible height error with 250mm bar length will produce 0.12  $\mu$ rad error. If this error is combined with several other errors, it will impact the final nano-accuracy.

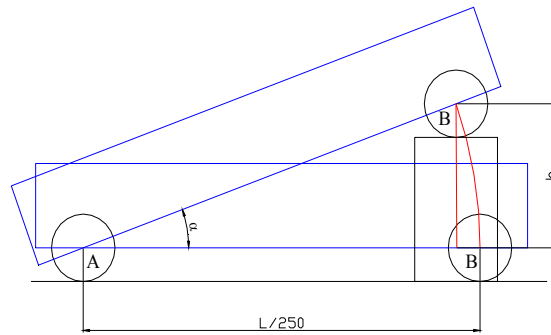


Fig. 17. Sine bar calibration scheme:  $\alpha = \arcsin (h/L)$

A computer controlled height lifting device and nanometer accuracy encoder are desired in order to reach nano-accuracy and to maintain temperature stability during the operation. The accuracy of mechanical and contacted Heidenhain length gauge is 30 nm. But its contact characteristic and insufficient accuracy will restrict nano-accuracy development of the sine bar system.

### 5.1.2 Tangent bar

Another relative angle metrology method uses the tangent function by means of two precise fixed distance sensors to measure the bar's tilt angle:

$$\alpha = \arctan (h / L) \quad (4)$$

where the h is the difference between the two distance-sensor readings and L is the distance between the two sensors (Fig. 18).

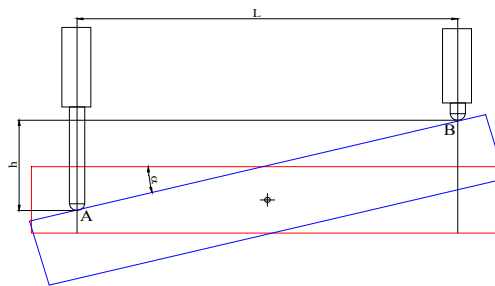


Fig. 18. Tangent bar calibration scheme:  $\alpha = \arctan (h/L)$

A vertical angle comparator (VAC) was developed at the BESSY-II Optics laboratory for the characterization and calibration of angle measuring sensors (Siewert et al., 2010). The VAC applies the tangent bar principle with a 1.3m long aluminum bar. The axis of the instrument is formed by a structure of a crossing flexure joint. A linear stepper motor actuator and two

linear encoders (Heidenhain CertoCP 60 K) enable a controlled tilting of the VAC. The linear encoder provides a constant uncertainty of  $\pm 50\text{nm}$  over a range of 60mm. The measuring resolution of the Certo is  $\pm 5\text{ nm}$ ; this corresponds to a tilting resolution of about 10 nrad. The achievable angular resolution of the VAC is about  $0.015\ \mu\text{rad}$ , limited by the performance of the linear stepper motor. The error budget of the VAC is estimated to 50nrad rms.

Recent developments in precise interferometric distance sensors has improved measurement accuracy to the one nanometer level, which significantly increases the reliability for trigonometric function calibration. The displacement interferometer can be used for small angle metrology. Fig. 19 is the example that uses a ZYGO system (ZYGO, 2003). The difference of two distance sensor readings divided by the distance between two points is the tilt angle. Its advantages are non-contact and nanometer accuracy. The National Research Council Canada uses this principle in their angle calibration systems (Pekelsky & Munro, 2005).

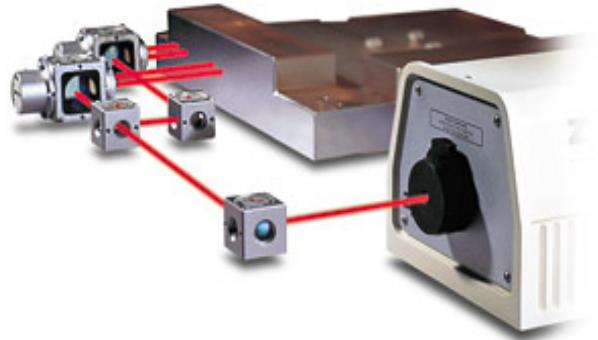


Fig. 19. Displacement Measuring Interferometers,  
Courtesy of ZYGO, <http://www.zygo.com/?/met/markets/stageposition/zmi/>

The fiber-based Fabry-Perot interferometer (FFPI) is a typical multiple-beam interferometer that can be used as a non-contact distance sensor. The space separating the reflecting surface is called the cavity length. The reflected light in the FFPI is wavelength modulated in exact accordance with the cavity length (Pullteap, 2010). The attocube's ultra-high precision displacement FPSensor has a repeatability of 1 nm at 20 mm cavity length, making it is suitable for small angle calibration with sine bar or tangent bar due to the advantages of high accuracy and non contact measurement characteristic (Attocube, 2010). Assuming its accuracy is 3 nm and it is used with a 250 mm sine bar, the error will be  $0.012\ \mu\text{rad}$ . In addition, its very compact size is very attractive. However, larger test angles will impact the test accuracy significantly.

## 5.2 Calibration by use of commercial angle testing devices

Instrument calibration with nano-accuracy can be done by use of commercial angle measuring instruments such as the theodolite, goniometer and other angle measurement

devices. However, they provide nanometer resolution or nanometer repeatability but rarely are nano-accuracy. They can not be used for nano-accuracy calibration over large angular ranges. There are only a few angle calibration devices that have been developed that can reach high accuracy with a large angular test range, which is expected by surface profiler.

The first calibration of the LTP angle error was made at ELETTRA in Italy in 1995 by use of a precision theodolite, Leica Wild T3000, with a sensitivity of  $0.1''$  (Qian et al., 2000). A small mirror M is fixed on the theodolite telescope in order to reflect the beam back to the LTP (Fig. 20). In this way we can know the mirror rotation angle precisely. The LTP records a stability scan as a function of time while the theodolite is rotated step by step with a separation of  $0.1^\circ$  or  $0.05^\circ$ . A step-like slope file can be obtained. The differences between the LTP angle values and theodolite reading angles determine the calibration angle error. By changing the LTP scale factor coefficient,  $e$ , the LTP angles can be adjusted. This is an absolute angle calibration. This test should be done after the precise adjustment of focal plane position has been done.  $0.1$  arc second ( $0.5 \mu\text{rad}$ ) was suitable for  $1 \mu\text{rad}$  accuracy calibration at that time but it is not sufficient for recent nano-accuracy calibration requirements unless a higher accuracy theodolite is available.

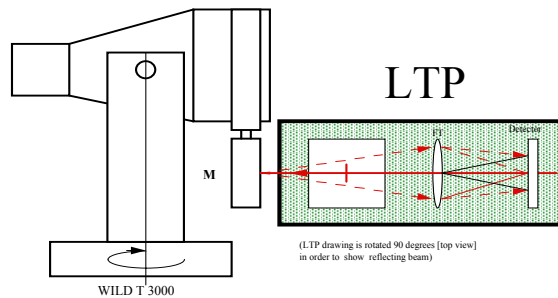


Fig. 20. Setup of precise LTP angular calibration: WILD T3000 is used as an accurate angle generator. While The LTP is making stability scan and data acquisition continuously, the theodolite angle is changed step by step

The PTB angle comparator is the most accurate standard angular measuring device today with a test range of  $360$  degrees (Probst et al., 1998). It is well suited for pencil beam profiler calibration, especially for the larger slope profiler test. The angle-measuring system of the comparator consists of a ring-shaped index disc of glass with a radial reflected-light phase grating with  $2^{17} = 131\,072$  graduation periods on a circle approx.  $400$  mm in diameter. Eight scanning heads uniformly distributed over the circumference of the graduation are used for scanning this graduation.  $2^{18} = 262\,144$  signal periods are formed in each scanning head, which corresponds to an angular period of approximately  $5''$ . Digital interpolation of the signal period with the factor  $2^{12} = 4096$  finally furnishes  $2^{30} = 1\,073\,741\,824$  measurement steps per  $360^\circ$ , which corresponds to an angle-measuring step of approx.  $0.0012''$  per scanning head. The angle value measured is finally obtained by averaging over all scanning heads. At present, this angle comparator allows an uncertainty of measurement of  $0.005''$  ( $k = 2$ ) to be reached.



Angular calibrations of profilers can be made at national standards bureaus. However, for researchers involved in precision angular R&D projects, it is necessary to check the nano-radian accuracy of a profiler frequently, and for calibration at remote sites, a low cost precision angle calibration device is desired.

In the case of sine bar and tangent bar systems for large angle measurements, the contact surface in using mechanical length gauge or reflection mirror surface in interferometric distance sensor is tilted. This will degrade the measurement accuracy considerably.

## 6. Precision metrology of in-situ and at-wavelength

Most measurements for precision optics are made in controlled environment in order to verify compliance with specifications. However, the actual in-situ use of the optical components could be very different from the laboratory condition. Beam quality and position can be affected by temperature instability, distortion under high vacuum condition, a noisy vibration environment, and thermal distortion due to absorption of high power beams from synchrotron and FEL sources. In these cases, on-line figure measurement of bending mirrors and adaptive optics is highly desirable. The manufacturing of large astronomical optical components with nanometer accuracy requires in-situ testing without removing the optics from the polishing machine. These in-situ situations present challenges to the metrology. Most of these metrologies are having increasing demands on nanometer level.

Owing to the restricted conditions for measurements on in-situ optics, very few measurements have been made, even though they are very important.

### 6.1 In-situ X-ray mirror thermal distortion measurement

Mirror distortion under high heat load has been recognized as a serious problem for third-generation synchrotron light sources, as well as for the first- and second-generation sources operating under conditions of reduced emittance, high current and with novel insertion devices. Efforts have been made at reducing mirror distortion by the use of high thermal conductivity or low expansion materials, cryogenic cooling, enhanced heat exchangers, jet cooling, and other means. Model calculations based on finite element design codes are used to predict distortion theoretically. However, the performance of any mirror will ultimately be determined experimentally when fully illuminated by the high power synchrotron beam, because the practical boundary conditions usually differ from the ideal theoretical conditions. An in-situ distortion test is then very useful for confirming the theoretical calculations. A precise measuring method to detect the in-situ distortion profile of a high heat load mirror for synchrotron radiation by use of the penta-prism Long Trace Profiler was developed in 1995 in Italy. A schematic diagram of the heat-load measurement equipment of the *in situ* LTP is shown in Fig. 21 (Qian et al., 1997, 1995). The optical head of the LTP II profiler is mounted horizontally on an optical table (TB). The first penta-prism (PT1) scans the sampling beam (SB) along the mirror under test (MUT) by use of a mechanical translation slide (MS) fixed to TB with a 250-mm travel length. The reference beam (RB) is directed onto a fixed spot on the MUT by a mirror (M1) and another penta-prism (PT2). This spot is located along the center of the length of the mirror but displaced by

15 mm transversely toward the edge. For symmetry reasons this point should not have a tangential slope variation component even if the mirror is subject to a high heat load.

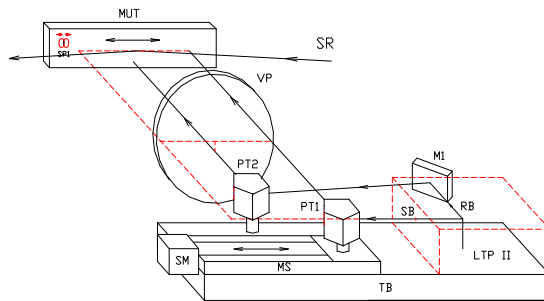


Fig. 21. Schematic of the in situ LTP test on beam line

A maximum distortion of  $0.47 \mu\text{m}$  over a length of 180 mm was measured for an internally water-cooled mirror on an undulator beam line at ELETTRA while exposed to a total emitted power of 600 watts (Fig. 22). For this measurement, the configuration with all of the equipment external to the vacuum chamber was used. The experiment has an accuracy and repeatability of 40 nm.

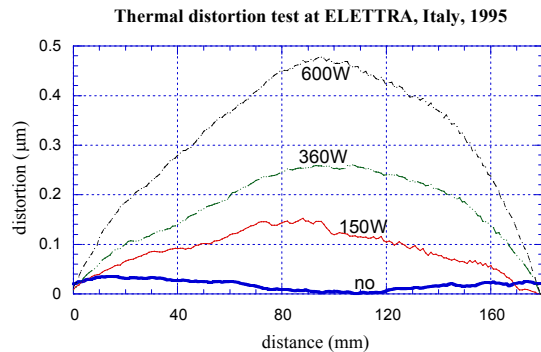


Fig. 22. In situ height distortion profiles of a synchrotron radiation mirror under a high heat load, as measured with the ppLTP: (a) Total power, 600W; undulator gap, 30mm; current, 181mA; energy, 2Gev. (b) Total power, 360W; undulator gap, 40mm; current, 187mA; energy, 2Gev. (c) Total power, 150W; undulator gap, 60mm; current, 224mA; energy, 2Gev. (d) No synchrotron beam on, thus corresponding to the test repeatability ( $<0.04\mu\text{m}$ , peak to valley)

The second thermal distortion test was done at The Advanced Photon Source (APS) at Argonne National Laboratory on the second mirror of SRI-CAT 2-ID-C beam line in 1997 (Takacs et al., 1998). The in-situ LTP scanned the central 90 mm of the 200 mm long mirror through a vacuum window while the mirror was subjected to heat loading from the

synchrotron beam. We measured about 200nm distortion in 90 mm when the X-ray beam was switched on. This distortion was not predicted by the finite element (FE) thermal calculations. By use of adjustable and movable slits upstream of the beam line, the illuminating beam spot position on mirror could be shifted. The aperture width was set to 0.5 mm, which corresponds to a 25 mm beam width on the surface. The very attractive feature of this test is that we can find the thermal bump produced by SR beam with great sensitivity (Fig. 23). The 200nm bump with a 20  $\mu$ rad slope followed the slits center shift movement. The results of this preliminary test indicate that after improvements in instrument stability, it should be able to measure 10nm thermal distortion with ease. This can be very powerful tool for nanometer spot mirror distortion test.

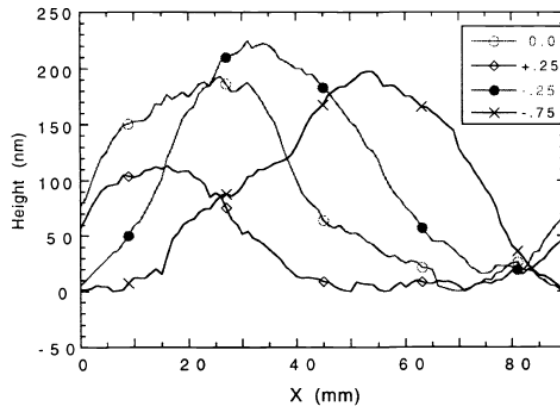


Fig. 23. Thermal bump produced on surface with 98mA beam current. Adjustable aperture is set at 0.5 mm H x 4.0 mm V, corresponding to a 25 mm wide beam on the surface. The aperture center was offset from the normal 0.0 mm position and the measurements show the corresponding lateral shift of the thermal bump. Unit in legend are mm.

## 6.2 In-situ precision metrology at optical workshop

In 2001, the portable LTP (PTLTP) was proposed for in-situ measurement at optical workshop to test optics during the final polishing step (Fig. 24) (Qian & Takacs, 2001). The portable LTP has a stationary optical head fixed to polishing machine. The optics to be polished, for example a cylinder mirror, is scanned by a moving table in order to scan the test beam over the mirror. A small mirror is fixed on the moving table in order to test the moving pitch error through a penta-prism.

An actual on-machine measurement has been made on measuring the slope and form errors of long cylindrical mirrors with optical tolerance precision and accuracy at RIKEN in Japan (Moriyasu et al., 2004). Cylindrical surfaces can be measured completely by steering the beam up the sides of the cylinder with a rotational mirror. To measure complete cylindrical surfaces, the direction of the beam to the measured object can be steered by controlling the angle of the flat mirror on the rotation table. In this way, a 2-D surface profile can be obtained.

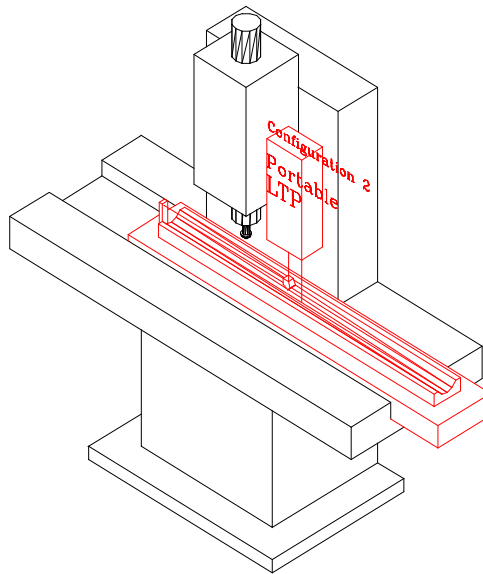


Fig. 24. Proposed in situ test at workshop in 2001

### 6.3 At wavelength metrology

Traditional optical metrology generally employs visible light as the illumination source, which is convenient and cost-effective. Recent developments in optics for extreme ultraviolet (EUV) lithography, x-ray synchrotron radiation optics, and free electron lasers operated at EUV and x-ray wavelengths place extraordinary requirements on the visible light metrology systems. Speer and collaborators (Harris et al., 1982; Speer et al., 1980) employed the point diffraction interferometer to perform at-wavelength measurements on grazing incidence optics into the far ultraviolet region. Sommargren, Goldberg and collaborators (Meddecki et al., 1996; Tejnli et al., 1997; Sommargren 1996; Goldberg et al., 1995) extended the technique into the EUV region to test zone plate wave-fronts and Schwarzschild optics quality in the 13nm wavelength region. This work was driven by the need to produce multilayer-coated normal incidence optics that were diffraction-limited at wavelengths to be used in soft x-ray projection lithography. A 0.5nm wave-front error is much more easily seen with a 13nm source, where it is  $\lambda/26$ , than with a 633nm HeNe source, where it is more than  $\lambda/1000$ .

The point diffraction interferometer requires a system that produces a point focus somewhere along the optical axis. A more general method for testing x-ray wave-fronts was developed by Weitkamp and collaborators (Weitkamp et al., 2005) utilizing Talbot effect Moire fringes generated by phase and absorption gratings. The interferometer can be placed anywhere downstream of an optical element where it is convenient, making it a very versatile technique. Wave-front distortions in the range of  $\lambda/100$  can be measured, where  $\lambda=0.1$  nm, and surface slope errors can be detected with an accuracy of better than 100 nrad over spatial periods from 1mm to 1 m.

Another versatile at-wavelength technique has been developed by Souvorov (Souvorov et al., 2002) and by Yumoto (Yumoto et al., 2006) and their coworkers, based upon the phase retrieval algorithms of Fienup (Fienup, 1982). By measuring the intensity variations in propagating x-ray beams downstream from a reflection or transmission element, the phase of the wave-front can be computed and projected back to the surface of the optical element. Yumoto has shown that surface figure errors measured by visible light interferometric means and by the phase retrieval methods agree to better than 1.5 nm on an 80mm long mirror. This method has the advantage over visible light interferometry in that it is sensitive to errors in the multilayer coatings on the optics, resulting in a measurement of the actual performance of the optic rather than just the profile of the top-most surface layer.

## 7. Thermal stability

### 7.1 Temperature stability requirement of $\pm 0.01^\circ\text{C}$ to ensure the $0.1\mu\text{rad}$ rms accuracy

After much experience with LTP stability scan measurements, we have seen that thermal variation has a significant impact on the ability to achieve  $0.1\mu\text{rad}$  rms slope error measurements. One can clearly see that slope error variations follow temperature fluctuations. Some precision profiler measurements require averaging of several tens of repeated measurements in order to achieve  $0.1\mu\text{rad}$  accuracy, and 2-D testing also requires much longer scanning time for multiple lines measurements in X and Y directions. In addition, the thermal drift effects are always delayed for one hour or more. So we use a 15 hour stability scan test as the standard procedure to verify measurement stability.

Fig. 25 is a stability comparison measurement between the LTP II and the PTLTP made at NSRRC in Taiwan (Qian & Wang, 2005). Test beams are sent from both LTP II and PTLTP to the same test point on the mirror through comparison unit, and then each reflected beam is divided into two beams back to two LTPs individually.

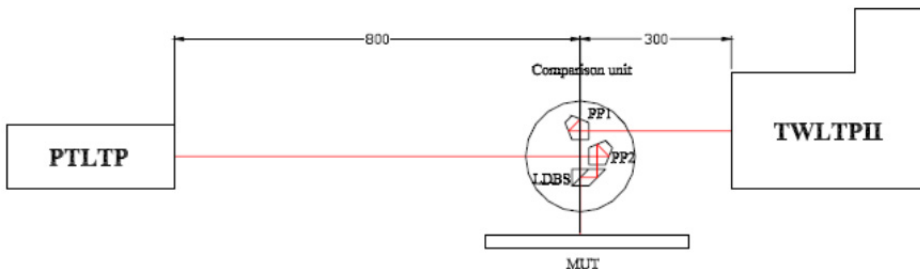


Fig. 25. Stability comparison test between the LTP II and the PTLTP made at NSRRC in Taiwan

Four files of stability scan are obtained simultaneously. Each LTP receives two reflected beams sent from itself and from the other. Fig.26 a) is the temperature oscillation during the test. Fig. 26 b) is the slope errors of 4 stability scans. 2 large slope error curves are: beam is sent from LTP II; and 2 small slope error curves are: beam is sent from PTLTP no matter it is accepted by which LTP. These indicate: a) Slope error exactly follows temperature fluctuation period but with some delay; b) The LTP II has large stability error due to adopting unstable non-monolithic beam splitting structure; c) When temperature oscillation

dropped to  $\pm 0.01^\circ\text{C}$  (flat section of the temperature curve), the stability over 15 hours was below  $0.1\mu\text{rad rms}$ .

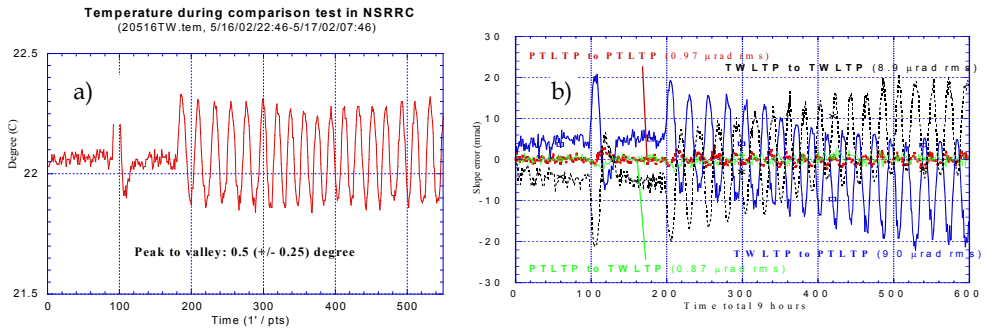


Fig. 26. a) LTPs temperature during the stability scan (without enclosure);  
 b) Stability comparison test: 1) 2 large slope error curves: beam is sent from LTP II;  
 2) 2 small slope error curves: beam is sent from PTLTP

Another stability measurement with an  $f=150\text{mm}$  LTP had a larger slope of  $0.17\mu\text{rad rms}$  due to larger temperature variation of  $0.07^\circ\text{C}$  (Fig. 27 a). The Fig. 27 b) shows the  $0.156\mu\text{rad rms}$  stability test of 15 hours in X-direction with an improved 2D CCD system ( $f'=400\text{ mm}$ ) while temperature variation of  $0.2^\circ\text{C}$  (P-V). These stability measurements verify that the  $\pm 0.01^\circ\text{C}$  temperature stability is necessary to ensure much less stability slope error than  $0.1\mu\text{rad rms}$ .

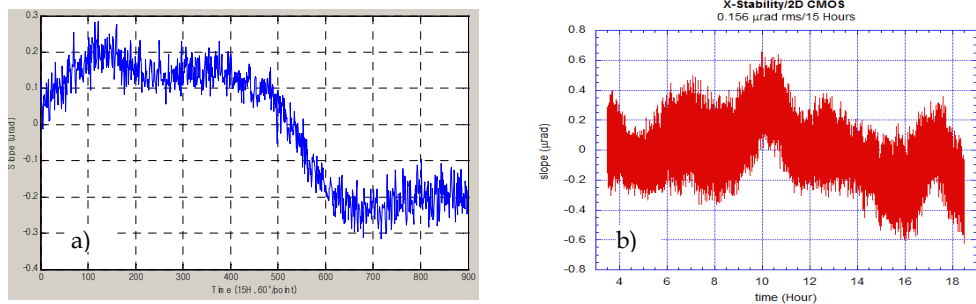


Fig. 27. a) Stability tested over 15 hours on  $f=150\text{mm}$  LTP with temperature variation of  $0.07^\circ\text{C}$ ; b) The 15 hours stability test in X-direction with an improved 2D system ( $f'=400\text{ mm}$ ) while temperature variation of  $0.2^\circ\text{C}$  (P-V)

### 7.2 Error reduction caused by thermal or mechanical drift

Recently, Yashchuk has developed measurement techniques that correct for slow thermal and mechanical drift errors that are inherent in typical LTP measurements (Yashchuk, 2009). His methods involve making a series of measurements on a surface in the forward and

reverse directions with the mirror rotated to the 180° and 0° orientations. Correcting these slow drift effects greatly improves the accuracy of the measurement. A project is underway to design a next-generation optical profiler that will incorporate an automated mirror rotation mechanism into the system so as to avoid operator intervention between scans (Yashchuk, 2011).

### 7.3 Temperature stabilization: arrange unavoidable thermal sources in enclosure only

It is better to keep unstable thermal sources outside the enclosure of the profiler as much as possible. The best choice: No drive motor is placed inside the enclosure.

It is necessary to apply monolithic wave-front splitting beam splitter (WSBS) to maintain good thermal stability.

Following is temperature variations of light source and CCD camera for reference:

- When USB is plugged in, the temperature of the CCD control board increases by 6.5°C; and the temperature of the CCD chip board varies by about 2.5°C (Fig. 28 a). However, both stabilize after about 30 minutes. The CCD chip board has to be set inside optical head in order to record the images, but it is possible to set the control board outside the optical head. The CCD should be turned on all the time in order to keep the temperature stable.
- Diode laser has very low power consumption, but it is still a considerable thermal source. Its temperature rise is about 1°C after turn on (Fig. 28 b). Again, it is better to place the diode laser outside of the enclosure and keep it switched on all the time for stabilization.

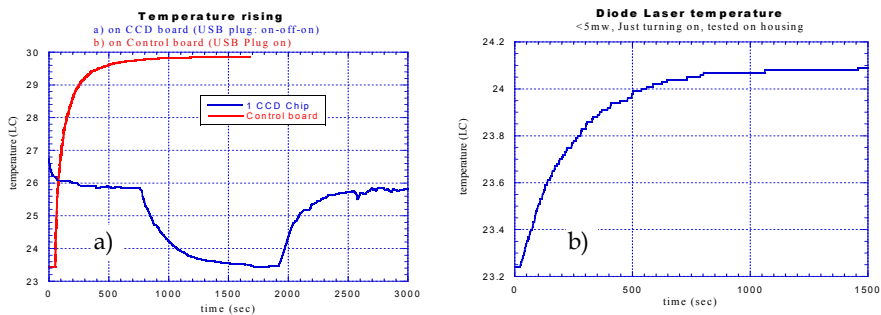


Fig. 28. a) Temperature variations of CCD board and control board; b) Temperature rising of diode laser

## 8. Disclaimer

Certain commercial equipment, instruments, or materials are identified in this document. Such identification does not imply recommendation or endorsement by the US Department of Energy, BNL, nor does it imply that the products identified are necessarily the best available for the purpose.

## 9. Acknowledgment

Notice: This manuscript has been authored by Brookhaven Science Associates, LLC under Contract No. DE-AC02-98CH10886 with the U.S. Department of Energy. The United States Government retains, and the publisher, by accepting the article for publication, acknowledges, a world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes.

## 10. References

- Alcock, S., K. Sawhney, K., & Scott, S. (2010). The Diamond-NOM: A non-contact profiler capable of characterizing optical figure error with sub-nanometre repeatability, *Nuclear Instruments and Methods in Physics Research, A* 616, 224-228
- Assoufid, L., Kim, K., Macrander, A. T., Lindberg, R., Kewish C. M., & Chubar, O., *Requirements for Grazing-Incidence Mirrors for Hard X-ray Free-Electron Laser Oscillator*, <http://www.esrf.eu/events/conferences/SMEXOS/PosterAssoufid>
- Assoufid, L., Bray, M., Shu, D. (2004). Development of a Linear Stitching Interferometric System for Evaluation of Very Large X-ray Synchrotron Radiation Substrates and Mirrors, *Eighth International Conference on Synchrotron Radiation Instrumentation. AIP Conference Proceedings*, V.705, 851-854.
- Attocube systems, <http://www.attocube.com/attoMETROLOGY/attoFPSensor.htm>, 2010
- Bhushan, B., Wyant, J.C., & Koliopoulos, C.L. (1985). Measurement of surface topography of magnetic tapes by Mirau interferometry, *Applied Optics* 24(10), p. 1489-1497.
- Becker, K. & Heynacher, E. (1987). M400 - A coordinate measuring machine with 10 nm resolution, in *In-process Optical Metrology for Precision Machining*, *Proc. SPIE* 802. pp. 209-216, presented at SPIE Symposium, The Hague, Netherlands,
- Corning HPFS® glass, 2008.  
[http://www.corning.com/specialtymaterials/products\\_capabilities/HPFS.aspx](http://www.corning.com/specialtymaterials/products_capabilities/HPFS.aspx)
- DeCew, A.E., Jr. & Wagner, R.W. (1986). An optical lever for the metrology of grazing incidence optics, in *Optical Manufacturing, Testing, and Aspheric Optics*, *Proc. SPIE* 645. pp. 127-132, Orlando, FL, 1-2 April 1986.
- Ennos, A.E. & Virdee, M.S. (1982). High accuracy profile measurement of quasi-conical mirror surfaces by laser autocollimation, *Precision Engineering* 4, p. 5-9.
- Engineering Synthesis Design Inc. (ESDI), Simultaneous Phase-Shifting for Vibration Intensive & Turbulent Environments, 2011  
[http://www.engsynthesis.com/pdfs/ESDI\\_intellium\\_H2000\\_DS.pdf](http://www.engsynthesis.com/pdfs/ESDI_intellium_H2000_DS.pdf)
- Fleig, J., Dumas, P., Murphy, P.E., and Forbes, G.W. (2003). An automated sub-aperture stitching interferometer workstation for spherical and aspherical surfaces, in *Proc SPIE* 5188, eds. A. Duparre and B. Singh. pp. 296-30
- Fienup, J.R. (1982). Phase Retrieval Algorithms - a Comparison, *Applied Optics* 21(15), p. 2758-2769.
- Goldberg, K.A., et al. (1995). At-Wavelength Testing of Optics for EUV, *Proc. SPIE* 2437, pp. 347-354.
- Heynacher, E. and Reinhardt, D. (1979). Measuring equipment for testing the directrix of high-resolution Wolter-type telescopes, *Proc. SPIE* 184, p. 167-169



- Harris, W., Mrowka, S., and Speer, R.J. (1982). Linnik Interferometer - Its Use At Short Wavelengths, *Applied Optics* 21(7), p. 1155.
- Irick, S.C., McKinney, W.R., Lunt, D.L., & Takacs, P.Z. (1992). Using a straightness reference in obtaining more accurate surface profiles from a long trace profiler for synchrotron optics, *Review of Scientific Instruments* 63(1), p. 1436-1438.
- Koliopoulos, C.L. and Wyant, J.C. (1980). Profilometer for Diamond-Turned Optics Using a Phase-Shifting Interferometer, *Journal of the Optical Society of America* 70(12), p. 1591.
- Kim, K., An X-Ray FEL Oscillator: Promises and Challenges, *Workshop on Physics and Applications of High Brightness Electron Beams*, November 16-19, 2009
- Kwang-Je Kim, An X-Ray FEL Oscillator: Promises and Challenges, *Workshop on Physics and Applications of High Brightness Electron Beams*, November 16-19, 2009
- Liu, Y.M., Lawrence, G.N., and Koliopoulos, C.L. (1988). Sub-aperture testing of aspheres with annular zones, *Applied Optics* 27(21), p. 4504-4513.
- Lammert, H., Siewert, F., & Zeschke, T. (2006). The Nano-Optic-Measuring Machine NOM at BESSY - further improvement of the measuring accuracy, in *BESSY Annual Report 2005*. BESSY GmbH: Berlin, Germany. p. 481-486.
- Li, H., Li, X., & Grindel, M.W. (1996). Measurement of X-ray Telescope Mirrors Using A Vertical Scanning Long Trace Profiler, *Opt. Eng.* 35(2), p. 330-338.
- Li, H., Takacs, P.Z., & Oversluisen, T. (1997). Vertical scanning long trace profiler: a tool for metrology of x-ray mirrors, *Proc. SPIE* 3152. pp. 180-187.
- Lightmachinery, Fluid Jet Polishing, <http://www.lightmachinery.com/Fluid-Jet-Polishing.html>, 2011
- Malacara, D. (1992). *Optical shop testing*. 2nd ed., New York: Wiley. xviii, 773 pp.
- Medecki, H., Tejnrl, E., Goldberg, K.A., & Bokor, J. (1996). Phase-shifting point diffraction interferometer, *Optics Letters* 21(19), p. 1526-1528.
- Mimura, H., Morita, S., et al. (2008). Focusing mirror for x-ray free-electron lasers, *Review of Scientific Instruments* 79, 083104.
- Mimura, H., et al. (2005). Relative angle determinable stitching interferometry for hard x-ray reflective optics, *Review of Scientific Instruments* 76(4), p. 045102-6.
- Moriyasu, S., Takacs, P. Z. et al. (2004). On-machine Metrology with LTP (Long Trace Profiler), *Proc. of SPIE* Vol. 5180, p. 385-392
- Murphy, P., Fleig, J., and Forbes, G. (2006). Sub-aperture stitching interferometry for testing mild aspheres, *Proc SPIE* 6293, eds. E.L. Novak, W. Osten, and C. Gorecki. pp. 62930J:1-10.
- Murphy, P., et al. (2003). Stitching interferometry: a flexible solution for surface metrology, *Optics & Photonics News* 14(5).
- NASA SBIR and STTR 2011 Program Solicitations, <http://sbir.nasa.gov/SBIR/sbirsttr2011/solicitation/>
- NPL, Calibration of Autocollimators at the NPL, 2010. <http://www.npl.co.uk/engineering-measurements/dimensional/dimensional-measurements/products-and-services/calibration-of-autocollimators>
- Pekelsky, J. R. & Munro, L. E. (2005). Bootstrap calibration of an autocollimator, index table and sine bar ensemble for angle metrology, *Proc. of SPIE* 58790D, p1-17
- Polack, F., Thomasset, M., & Brochet, S. (2010). An LTP stitching procedure with compensation of instrument errors: Comparison of SOLEIL and ESRF results on

- strongly curved mirrors, *Nuclear Instruments and Methods in Physics Research Section A*, V.616, Issue 2-3, 207-211.
- PTB Working Group 8.42, Form measurement of curved optical surfaces, 2011  
<http://www.ptb.de/cms/index.php?id=formmessung-842&L=1>
- Pullteap, S. (2010). Development of a Fiber based Interferometric Sensor for Non-contact Displacement Measurement, *World Academy of Science, Engineering and Technology* 66, p.1164-1168
- Probst, R., Wittekopf, R., & Krause, M. (1998). The new PTB angle comparator, *Meas. Sci. Technol.* 9 1059 (1998) doi:10.1088/0957-0233/9/7/009
- Pruss, C., Garbusi E., & Osten, W., Testing Aspheres, *Optics and Photonics News*, Vol. 19, Issue 4, 24-29.
- Qian, S. N. (2011). Scanning Optical Head with Nontilted Reference Beam: Assuring Nanoradian Accuracy for a New Generation Surface Profiler in the Large-Slope Testing Range, Hindawi Publishing Corporation, *International Journal of Optics*, Volume 2011, Article ID 902158, 9 pages, doi:10.1155/2011/902158
- Qian, S. N. & Qian, K. (2010). Study and Considerations of Nanometer and Nano-radian Surface Profiler, *SPIE Proc.* Vol. 7656 76560D-76560D-10.
- Qian, S. N. & Takacs, P. (2007). Design of multiple-function long trace profiler, *Optical Engineering* 46(4), p. 9.
- Qian, S. N., Wang, D. (2005). Real-time stability and profile comparison measurements between two different LTPs, *Proc. SPIE* V5921, pages 59210K, doi:10.1117/12.618791
- Qian, S. N., Wang, Q., Hong, Y., and Takacs, P. (2005). Multiple functions long trace profiler (LTP-MF) for National Synchrotron Radiation Laboratory of China, *Proc. SPIE* 5921, p. 592104.
- Qian, S. N. & Takacs, P. Z. (2004). Equal optical path beam splitters by use of amplitude-splitting and wavefront-splitting methods for pencil beam interferometer, *Proceedings of SPIE* Vol. 5193, 79-88
- Qian, S. N., & Takacs, P. Z. (2003). Wave front-splitting phase shift beam splitter for pencil beam interferometer, *Review of Scientific Instruments*, Vol.74, No. 11, 4881-4884, Nov.
- Qian, S. N. & Takacs, P. Z. (2001). Portable Long Trace Profiler: Concept and solution, *Review of Scientific Instrumentations*, Vol.72, No. 8, 3198-3204, Aug.
- Qian, S. N., Sostero, G., & Takacs P. Z. (2000). Precision calibration and systematic error reduction in the long trace profiler, *Opt. Eng.* 39, 304; doi:10.1117/1.602364
- Qian, S. N., Jark W., et al. (1997). Precise measuring method for detection the in situ distortion profile of a high-heat-load mirror for Synchrotron radiation by use of a pentaprism long trace profiler, *Applied Optics*, Vol. 36, No. 16, 1 June.
- Qian, S. N., Jark, W., Takacs, P. Z. et al. (1995). In-situ surface profiler for high heat load mirror measurement, *Opt. Eng.*, vol. 34 (2), 396-402, Feb.
- Qian, S. N., Jark, W. & Takacs, P. Z., The penta-prism LTP: A long-trace-profiler with stationary optical head and moving penta prism, *Rev. Sci. Instrum.* 66, 2187 (1995);
- Soufli, R., Surface metrology and polishing techniques for current and future-generation EUVL, *International Workshop on EUV Lithography*, Hawaii June 16, 2011

- Schulz, M., Ehret, G., Stavridis, M., & Elster C. (2010). Concept, design and capability analysis of the new Deflectometric Flatness Reference at PTB, *Nuclear Instruments and Methods in Physics Research A* 616, 134–139
- Schulz, M., et al. (2010). Scanning deflectometric form measurement avoiding path-dependent angle measurement errors, *JEOS Rapid Publications* 5, 10026
- Schulz, M., Wiegmann, A., Márquez A., & Elster, C.(2008). *Optical flatness metrology: 40 years of progress, Opt. Pura Apl.* 41 (4) 325-331.
- Schott, TIE26: Striae in optical glass, 2004,  
[http://www.us.schott.com/advanced\\_optics/english/download/schott\\_tie-26\\_homogeneity\\_of\\_optical\\_glass\\_july\\_2004\\_us.pdf](http://www.us.schott.com/advanced_optics/english/download/schott_tie-26_homogeneity_of_optical_glass_july_2004_us.pdf)
- Siewert, F., Buchheim, J., & Zeschke, T. (2010). Characterization and calibration of 2<sup>nd</sup> generation slope measuring profiler, *Nuclear Instruments and Methods in Physics Research, A* 616, 119-127
- Siewert, F., et al. (2004). The Nanometer Optical Component Measuring Machine: a New Sub-nm Topography Measuring Device for X-ray Optics at BESSY, *AIP Conference Proceedings* 705(1), p. 847-850.
- Souvorov, A., et al. (2002). Deterministic retrieval of surface waviness by means of topography with coherent X-rays, *Journal of Synchrotron Radiation* 9, p. 223-228.
- Sommargren, G.E. (1996). Phase shifting diffraction interferometry for measuring extreme ultraviolet optics, *OSA Trends in Optics and Photonics. Vol.4 Extreme Ultraviolet Lithography*, p. 108-112.
- Stedman, M. & Stanley, V.W. (1979). Machine for the Rapid and Accurate Measurement of Profile, in *Advances in Optical Production Techniques, Proc. SPIE* 163. pp. 99-102, Sira, London,
- Sarnik, A. and Glenn, P. (1987). Mirror figure characterization and analysis for the Advanced X-ray Astrophysics Facility/Technology Mirror Assembly (AXAF/TMA) x-ray telescope, *Proc. SPIE* 830, eds. S. Bowyer and J.C. Green. pp. 29-36
- Speer, R.J., et al. (1980). Soft-X-Ray Performance Of Toric Gratings Fabricated With Plane-Waves, *Nuclear Instruments & Methods* 172(1-2), p. 303-306.
- Takacs, P.Z., Church, E.L., Bresloff, C.J., and Assoufid, L. (1999). Improvements in the accuracy and repeatability of long trace profiler measurements, *Applied Optics* 38(25), p. 5468-5479.
- Takacs, P. Z., Qian, S. N., Randa, K. J., et al. (1998). Mirror Distortion Measurements with an In-Situ LTP, *Proc. SPIE* Vol. 3447, p. 117-124.
- Takacs, P.Z., Furenlid, & K., DeBiasse, R. (1989). Surface topography measurements over the 1 meter to 10 micrometer spatial period bandwidth, *Proc. SPIE* 1164, pp. 203-211.
- Takacs, P. Z. & Qian, S.N.,(1989). Surface profile interferometer, *United State Patent* No. 4,884,697 (5 December 1989).
- Takacs, P.Z., et al. (1988). Long trace profile measurements on cylindrical aspheres, *Proc. SPIE* 966, pp. 354-364
- Takacs, P. Z., Qian, S. N., & Colbert, J. (1987). Design of a long-trace surface profiler, in *Metrology-Figure and Finish, Proc. SPIE* 749, 59

- Takacs, P.Z. & Bresloff, C.J. (1986). Significant Improvements in Long Trace Profiler Measurement Performance, *Proc. SPIE* 2856, eds. L.E. Berman and J. Arthur. pp. 236-245.
- Tejnil, E., et al. (1997). At-wavelength interferometry for extreme ultraviolet lithography, *Journal of Vacuum Science & Technology B* 15(6), p. 2455-2461.
- von Bieren, K. (1982). Pencil Beam Interferometer for Aspherical Optical Surfaces, *Proc. SPIE* 343, pp. 101-108.
- von Bieren, K. (1983). Interferometry of Wavefronts Reflected Off Conical Surfaces, *Appl. Opt.* 22, p. 2109-2114.
- Wang, Q., Guangjun Gao, G., & Griesmann, U., Radius Measurement of Spherical Surfaces With Large Radii-of-Curvature Using Dual-Focus Zone Plates, *Conference of Optical Fabrication and Testing*, October 19, 2008, [http://www.nist.gov/customcf/get\\_pdf.cfm?pub\\_id=824680](http://www.nist.gov/customcf/get_pdf.cfm?pub_id=824680)
- Weitkamp, T., et al. (2005). X-ray wavefront analysis and optics characterization with a grating interferometer, *Applied Physics Letters* 86(5), p. 054101.
- Wiegmann, M., Schulz, M., & Elster C. (2010). Improving the lateral resolution of a multi-sensor profile measurement method by non-equidistant sensor spacing, *Opt. Express* 18, 15807-15819
- Wyant, J.C., Koliopoulos, C.L., Bhushan, B., and George, O.E. (1984). An optical profilometer for surface characterization of magnetic media, *ASLE Transactions* 27(2), p. 101-113.
- Yamauchi, K., et al. (2011). Single-nanometer focusing of hard x-rays by Kirkpatrick-Baez mirrors, *J. Phys.: Condens. Matter* 23 394206
- Yamauchi, K., et al. (2003). Microstitching interferometry for x-ray reflective optics, *Review of Scientific Instruments* 74(5), p. 2894-2898.
- Yashchuk, V.V., et al. (2011). Development of a new generation of optical slope measuring profiler, *NIM A* 649(1), p. 153-155.
- Yashchuk, V.V. (2009). Optimal measurement strategies for effective suppression of drift errors, *Review of Scientific Instruments* 80(11), p. 115101-10.
- Ye, H., & Liming Yang, I. (2011). Accuracy and analysis of long-radius measurement with long trace profiler, *Chinese Optics Letters*, Vol. 9, Issue 10, pp. 102301
- Youngworth, R. N., & Stone, B. D. (2000). Simple estimates for the effects of mid-spatial-frequency surface errors on image quality, *Appl Opt.* May 1; 39(13):2198-209.
- Youngworth, R. N., DeGroot, J. E., & D. M. Aikens, D. M. (2008) Specification and Control of Mid-Spatial Frequency Wavefront Errors in Optical Systems, *conference paper, Optical Fabrication and Testing (OFT)*, Rochester, NY, October 19, 2008
- Yumoto, H., et al. (2006). At-wavelength figure metrology of hard x-ray focusing mirrors, *Review of Scientific Instruments* 77(6), p. 063712.
- Yumoto, H., et al. (2010). Stitching-angle measurable microscopic-interferometer: Surface-figure metrology tool for hard X-ray nanofocusing mirrors with large curvature, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 616(2-3), p. 203-206.
- Zecchino, M. Optical Metrology for Large Telescope Optics.pdf, Portions of this article appeared as *Dynamic Metrology Perfects Reflections*, Laser Focus World, February 2008.

[http://www.4dtechnology.com/reflib/Optical Metrology for Large Telescope Optics.pdf](http://www.4dtechnology.com/reflib/Optical%20Metrology%20for%20Large%20Telescope%20Optics.pdf)

ZYGO, Stage Position Metrology, ZMI™ Series Displacement Measuring Interferometers

<http://www.zygo.com/?/met/markets/stageposition/zmi/>

ZYGO, Displacement-measuring interferometers provide precise metrology, *Laser focus world*, 12/01/2003

## **Section 3**

### **Voltage, Current and Frequency**



# Josephson Junctions for Present and Next Generation Voltage Metrology

Vincenzo Lacquaniti and Andrea Sosso  
*Istituto Nazionale di Ricerca Metrologica, Torino,  
Italy*

## 1. Introduction

Josephson array voltage standards are the most complex superconducting integrated circuits in actual use, and represent the first and the most outstanding achievement in quantum metrology.

However some issues must still be improved for a more extended exploitation of these devices in basic metrology and in advanced precision instrumentation for applications ranging from aerospace, to defense and others.

This is mainly true when the extension of DC voltage references based on Josephson effect to AC or to the generation of programmable voltages is considered, since the circuits of these arrays are usually more complex, requiring a higher number of junctions.

In this chapter we present the most developed and the most advanced solutions studied so far to satisfy this problem, which require to achieve the best integration and the lowest power dissipation, while preserving the basic issue of the highest noise immunity of the voltages produced by the arrays.

These characteristics should be met through the most suitable selection of some Josephson junctions parameters such as the maximum critical current density, the characteristic voltage, the junction size and the degree of hysteresis of the current voltage characteristic.

A nontrivial role is also played by the thermal stability of some of these parameters.

In order to obtain the best performances, and therefore to optimize the characteristics outlined above, we will examine different types of Josephson junctions, discussing the advantages and disadvantages of each type and showing possible research directions to choose the most suited fabrication materials and processes, considering also the possibility of using cryocoolers.

The contents of the chapter are the following:

In a first section, Josephson effect and Josephson junctions, some basic physical concepts are reviewed

In the second section, Josephson Voltage Standards: from DC to AC, the application of the quantum effect to Voltage metrology is outlined, emphasizing the specificity of the extension to AC respect to the traditional DC standard



The main focus of section 3, Current junction technologies and fabrication issues, is on the different type of Josephson junctions employed so far in circuits for programmable and AC voltage standard, with details on the link between fabrication technology and device performances.

In a last section, Special issues for next generation standard and possible solutions the present and future applications of these standards to measurement are mentioned, with main stress on the routes to overcome present limitations to these challenges

## 2. Josephson effect and Josephson junctions

Modern voltage metrology is based on the Josephson effect which is a macroscopic quantum effect of inherent fundamental precision occurring in a weak link between two superconductors.

The essential properties of superconductive material are that the resistance to a continuous electrical current becomes zero and a magnetic field cannot penetrate inside, up to a critical value, which is characteristic of each material. Superconductivity is therefore described by the conditions electrical resistance  $R=0$  and magnetic induction  $B=0$ . The superconductive, non dissipative state, is characterized by an advantageous energetic condition respect to the normal state, since in some materials below a critical temperature  $T_c$  and a critical applied magnetic field  $H_c$ , all the electrons condensate at the same energy, which is separated by the Fermi level from a gap  $\Delta$ , in the range of about 1 mV.

This energetically favorable superconductive state establishes a long range order extending through the material, where the charge carriers, the Cooper pairs, flow without dissipation. The Cooper pairs are a particle which has twice the electric charge of the electron and a virtual radius corresponding to the coherence length of the superconductor, of the order of tenths of micrometer. They are formed at low temperature by the mutual attraction occurring between two electrons exchanging virtual phonons related with vibrations of the lattice [1].

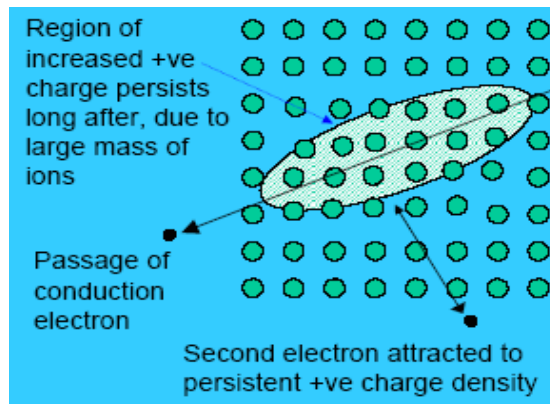


Fig. 1. Picture of the Cooper pair binding mechanism inside the atomic lattice.

Concerning the Josephson effect, it was firstly theoretically predicted and observed for tunnel junctions with two superconductive electrodes separated by a very thin insulating

barrier (of the order of nm) [2]. In this first model, for a Josephson junction realized as a weak link between two superconductors, a quasiparticle inside the barrier, in order to penetrate in one of the superconductors must get an energy greater than the superconductor energy gap, which is not possible by direct charge transfer.

But another process may occur: in the so called Andreev reflection [3] an electron impinging at the interface of the weak link with the superconductor, is converted into a hole moving in the opposite direction, thus creating a Cooper pair. Similarly, at the other side of the weak link, the reflection process annihilates a Cooper pair from the other superconductor. The result is the tunneling of the Cooper pairs through the barrier of the weak link. Nowadays, it was generalized to a variety of junctions, the most accepted interpretation is based on the Andreev reflection process.

This tunneling is not dissipative up to a critical value of current  $I_c$ , whose amplitude is determined by the geometry of the junction, the materials and the operating temperature. Below this value the current depends only from the phase difference  $\varphi$  of the two wave functions describing the superconductive electrodes of the junction.

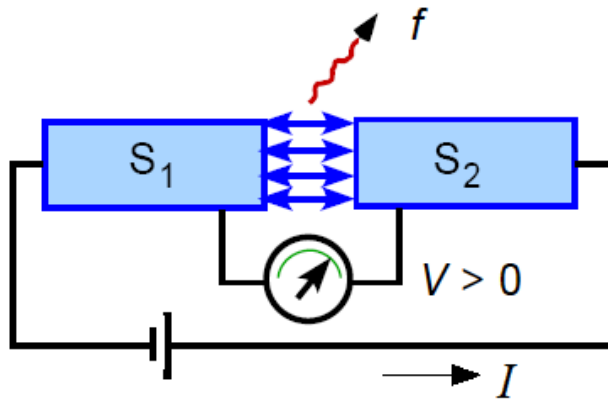


Fig. 2. Josephson junction general scheme. When a voltage drop  $V$  appears between the two superconductive electrodes an oscillating signal of frequency  $f$  is emitted by the junction.

The first Josephson equation, called DC Josephson effect then states:

$$I = I_c \sin \varphi \quad (1)$$

When the biasing current  $I$  is  $> I_c$ , a voltage  $V$  appears across the junction and this voltage generates on its own an oscillating current, whose frequency  $f$  is related to the voltage from the second Josephson equation.

$$V = (h/2e) \times f \quad (2)$$

The last relation is at the basis of the modern Voltage metrology. The quantity  $h/2e$  is called  $\Phi_0$ , and represents the unity flux quanta.

As stated, equation 2) is independent from any device characteristic or any operating condition, and does not have any drift in time.

In the practical experiment, the Josephson frequency, which is in the THz range, is modulated by a microwave signal, producing voltage steps at fixed, constant interval  $n (h/2e) \times f$ , with  $n$  integer 1, 2,.. These steps, called Shapiro steps, are used as reference voltage in the measurements, and their amplitude in current depends on the power of the microwave signal coupled to the junction, according with good approximation to the relationship: (3) These steps, which are what is used as reference voltage in the measurement, are called Shapiro steps, and their amplitude depends on the power of the microwave signal coupled to the junction.

$$I_s = I_c \sum_{n=0}^{\infty} (-1)^n J_n \left( \frac{V_1}{\Phi_0 f_1} \right) \sin \left[ \varphi_0 + \frac{2\pi}{\Phi_0} V_0 t - 2\pi n f_1 t \right] \quad (3)$$

where:  $I_s$  step current width,  $I_c$  critical current,  $I_s$  step current width,  $V_0$  ( $V_1$ ) dc (ac) voltage bias,  $f_1$  ac bias frequency, and  $J_n$  Bessel function of order  $n$ .

Experiments where the voltage provided by two different Josephson devices radiated by different microwave signals was detected through a Superconductive Quantum Interference Device, SQUID, the present most sensitive flux and current detector (resolution of the order of.  $\varphi_0 = 2.068 \times 10^{-15}$  Wb), showed no difference also in measurement repeated along the time, at the level of  $10^{-16}$ .

The device realizing the relations 1) and 2), the Josephson junction, can be obtained in different ways, according to what mentioned above. In particular, for the application to Voltage standard, after first point contact junctions, which were not reproducible, and generated voltages around 1 mV, thin film junctions have been used, exploiting the development of microelectronics processes.

The motion equation of the Josephson junction, describing the current flow is the electrical equivalent of the mechanical pendulum, and gives the total current in the junction as sum of three terms

$$I = I_c \sin \varphi + VR + C dV/dt, \quad (4)$$

where  $R$  and  $C$  are equivalent resistance and capacitance of the junction

Defining the parameter  $\beta c = (2\pi I_c R^2 C) / \varphi_0$  as the damping factor of the junction in analogy with the pendulum, the current voltage, I V characteristic of a Josephson junction can be single valued for  $\beta c < 1$ , and the junction is called non-hysteretic or overdamped, or multiple valued, for  $\beta c > 1$ , and the junction is called hysteretic or underdamped.

Correspondingly, the response to a microwave signal produce two type of staircase, as will be reported in the next section.

### 3. Voltage standards: From DC to AC

A DC voltage  $V$  applied to a Josephson junction equals exactly the rate of single flux quanta  $f\Phi_0$ , transferred along the junction:  $V = \Phi_0 f$ . If the flux transfer is phase-locked to an external oscillator with the highly stabilised frequency  $f$ , the transfer rate is kept constant over a certain range of current through the junction: this leads to constant voltage steps in the DC characteristic at  $V_n = n\Phi_0 f$ . Due to the nonlinearity of the DC characteristic, this occurs not only at the fundamental frequency  $f$  but also at the higher harmonics ( $n = 2, 3, 4, \dots$ ). For a

drive frequency of 70 GHz, the fundamental voltage step amounts to about 135  $\mu\text{V}$ . As the frequency can be stabilised to a very high degree, such a voltage step is a perfect voltage reference with a reproducibility of up to a few parts in  $10^{10}$ .

Josephson junctions used nowadays in DC voltage standard applications are based on hysteretic SIS junctions with zero crossing steps i.e. voltage steps whose current range spans positive and negative values, including the condition of zero DC bias. The choice of this technique dates back to the first attempts in series-connecting the thousands of junctions needed to reach output levels of 1 V and above, as required in metrological applications. The voltage steps are generated in the sub-gap part of the DC characteristic and all voltage steps between the maximum output voltages, e.g. -10 V and +10 V are biased by the same current. Exploitation of the zero crossing steps, first suggested by Levinsen [4] eventually allowed to overcome many technological difficulties and made it possible to realize arrays with reproducible overlapping steps.

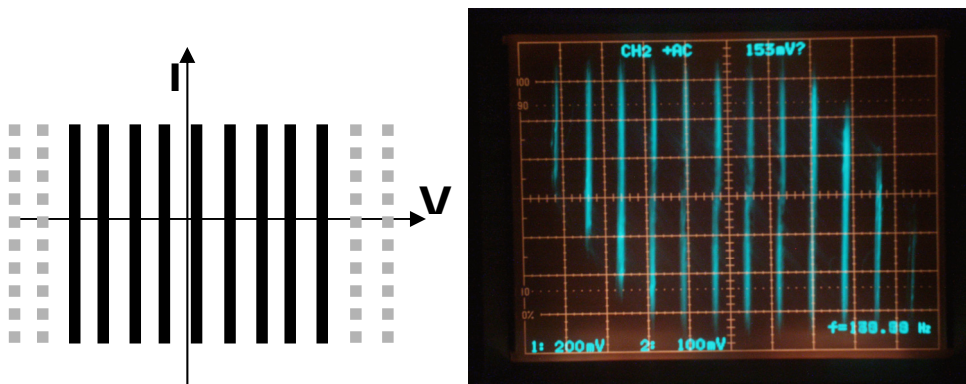


Fig. 3. Operating principle of programmable Josephson arrays (right) and IV characteristics (left) of the INRiM 10 V standard observed at the oscilloscope.

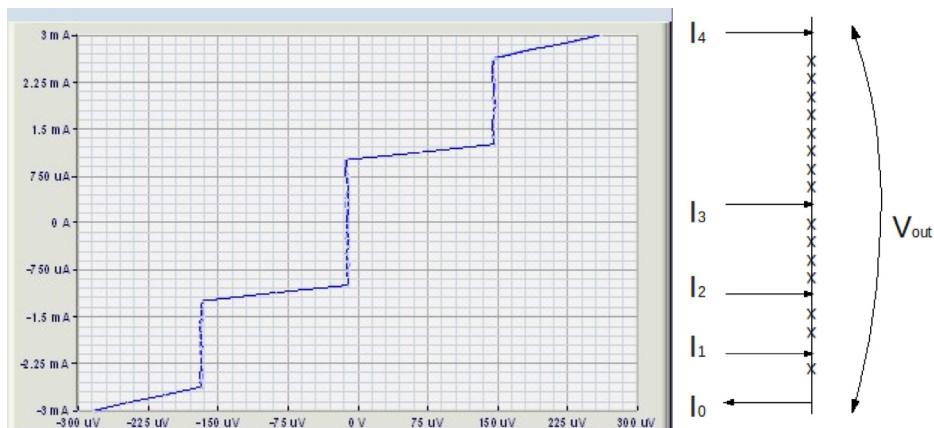


Fig. 4. Operating principle of programmable Josephson arrays. By controlling the bias currents  $I_1 - I_4$  it is possible to change the output voltage  $V_{\text{out}}$ .

To increase the output reference voltage to practical values, large series arrays of strongly under-damped Josephson tunnel junctions of the SIS-type (Superconductor-Insulator-Superconductor) - in this case Nb-Al<sub>2</sub>O<sub>3</sub>-Nb - with hysteretic DC characteristic were developed. On the one hand, this allows to make use of higher harmonics steps up to  $n = 6$  on an average per junction and generate 10 V output voltages with the relatively low number of 12000 to 14000 junctions. The availability of 10 V standards with quantum accuracy has led to dramatic improvements in DC voltage metrology, and it is now possible in primary DC voltage calibrations at 10 V to attain relative uncertainties as low as  $10^{-11}$ .

More recently, the interest in voltage standard research has moved to the investigation of techniques for extending the application of Josephson arrays to AC quantum standards and to standards for arbitrary time-varying signals. To this aim, junctions with non hysteretic behavior were suggested to allow changing the output voltage through control of the bias current. The substantial difference, from the application viewpoint, in using non hysteretic junctions, is that their IV curve (voltage *vs.* current relationship) under irradiation is a one to one staircase, thus the output voltage is univocally defined by the current feed through the bias circuit. This is not the case for hysteretic junctions used in DC standards, where steps are overlapping and all share approximately the same interval of currents.

In so-called programmable standards, the junctions bias currents are used to activate/deactivate array sections. Such arrays are typically subdivided in sub-circuits with series connected junctions generating voltages following a power of two sequence. Combining the sections it is then possible to source binary programmed voltages in a way that is very similar to the technique used in electronic digital to analog converters [5]. In order to replace best AC standards, the uttermost accuracy has to be reached and many efforts have been devoted in realizing arrays with performances suited to the tight requirements set by modern primary metrology. Many approaches to junction fabrication have been developed, and several different technologies have proven successful in providing voltages up to 10 V, with good metrological properties. Programmable Josephson arrays are so far the most successful attempt to extend metrological applications of Josephson standards beyond dc. Programmable arrays operating at 1 V have been effectively used for several applications: as traveling standards for international comparisons [6], for generating precisely varying voltages in a watt balance [7], as quantum impedance and power standards [8]. Moreover, only programmable standards can presently provide output voltages up and above 1 V, and even exceeding 10 V [9].

Nonetheless, this technique suffers from some limitations, the most severe is due to the time for step switching, where junctions are not operating in a quantized state. During these transients, the array voltage is not precisely known and the uncertainty of the generated signal increases with the fraction of period spent in the transients. Since the minimum transient time is constrained by technological limitations, programmable arrays can fulfill primary metrology uncertainties only for signals with frequencies up to few hundreds Hz [10].

To overcome limitations of programmable standards, arrays operating with a pulsed, square wave, rf signal have been developed. Using short pulses instead of a sinusoidal rf signal

makes it possible to effectively modulate the signal period while keeping junctions phase locked over a wide range of frequencies [11],[12],[13]. Fundamental accuracy follows from the control of the flux quanta transferred through the junctions by the pulsed signal. The output voltage is then exactly calculable in terms of fundamental constants if the number of the flux quanta per unit time, i.e. the pulse repetition rate is known [14]. Pulsed standards allow to synthesize arbitrary waveforms with quantum accuracy based on the sigma-delta technique for digital to analog conversion developed for semiconductor electronics and very high spectral purity [15]. Both operation and fabrication of pulsed standards set very challenging problems. Due to the complexity of pulse waveform, the apparatuses for generation of precisely frequency controlled pulses are sophisticated and expensive, and the design of rf transmission lines is extremely difficult because of the harmonic richness of the signal. In addition, it is extremely difficult to generate a bipolar output with a frequency modulated Josephson array, and very complex AC biasing techniques, involving a sine wave and a pulsed signal, both synchronized, must be used for real AC operation [16]. Generation of the pulse train for proper junction operation requires top-end instrumentation and unavoidably limits the signal fundamental frequency to values much lower than those obtainable with continuous wave sources. Power distribution to array junctions is also of concern, since the usual microwave techniques developed for nearly monochromatic signals are not directly applicable to broadband pulses. Adoption of lumped circuit methodologies seems at present the most viable solution, though highly demanding on the fabrication side. In order to have negligible effects on circuit behavior, the propagation time of rf signal along the array must be smaller than the signal period, i.e. the array dimensions must be smaller than the signal wavelength  $\lambda$ . To guarantee reliable operation,  $\lambda/8$  is typically considered the maximum value acceptable for array dimensions. Despite all these difficulties, arrays with as many as 10000 junctions have been successfully fabricated, and synthesized voltage signals up to 275 mV rms have been generated with quantum accuracy and extreme spectral purity [17].

Type of Voltage standard	Development status	Performances limitation and typical linewidths	Open questions	Future expected impact
DC voltage standard	1 and 10 V with part in $10^{10}$ accuracy, used world wide	10 $\mu\text{m}$		Present status: standard calibration tool
Programmable voltage standard	1V available, 10 V few samples, first applications	AC voltages < 3 KHz 5 $\mu\text{m}$	To reduce the influence of transients	World wide use for improved AC metrology
Pulse driven voltage standard	275 mV, pure spectra, first applications	1 $\mu\text{m}$ and less	Increase output voltage to 1 and 10 V	World wide use for improved AC metrology

Table 1. Comparison of different type of voltage standard.

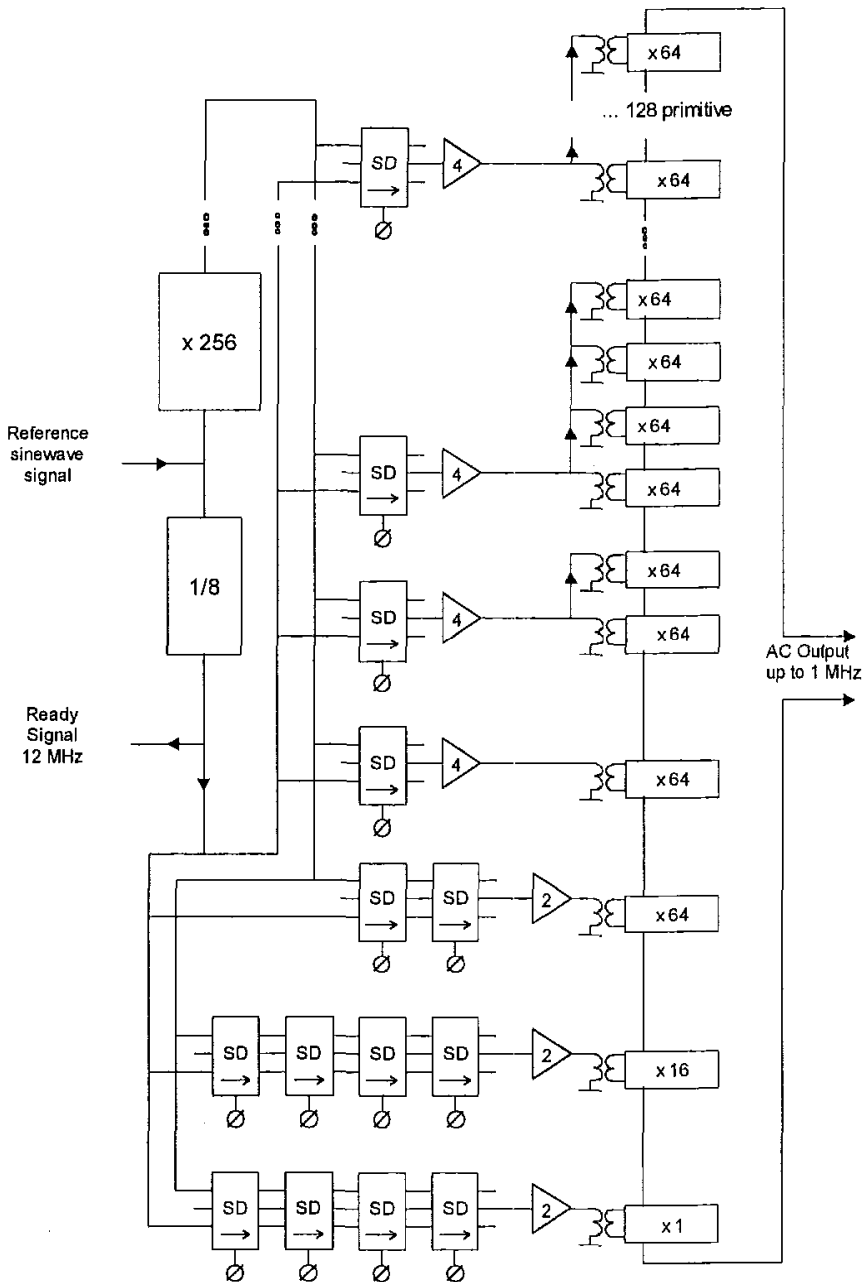


Fig. 5. Simplified schematic of a Digital to Analog converter realized with RSFQ technology (from [20]).

A new technique has been recently proposed to overcome some of the difficulties encountered with programmable and pulsed standards. The method, named Pulse Power Modulation (PPM), is based on a controlled activation/deactivation of the driving rf signal, to simultaneously set every junction in the array into either one of two states: zero voltage in absence of rf signal and a quantum defined step voltage upon rf application. Waveform synthesis can then be realized by Pulse Width Modulation of the array voltage [18]. The requirements in junction technology for PPM are different from those set by programmable and pulsed standards, since for proper operation an IV curve where the interval of currents of the relevant steps partially overlaps the critical current is needed. Such a behavior can be obtained if a precise control over junction hysteresis is feasible in the fabrication stage, to provide an intermediate degree of hysteresis, between those of the fully hysteretic SIS for DC and the non hysteretic SNS.

Although limited in space, this overview wouldn't be complete without mentioning RSFQ as an alternative technique to synthesize arbitrary and AC signals with quantum voltage accuracy. RSFQ has been, and still remains, a very active field of research for quantum digital electronics applications [19], yet developments for voltage standards have always been left out of the mainstream of research interests in Metrology. One of the main reasons for that is most likely to be found in the completely different approach, know-how and experiences involved by RSFQ, with respect to the common background of voltage metrologist. In RSFQ standards, accurate voltage signals are generated by controlling flux quantized by a Josephson junction. Flux quanta generation is "triggered" by a pulse sequence, thus can be precisely timed, in analogy with the "phase lock" process exploited in array standards [20]. RSFQ approach is potentially advantageous in that the complex and expensive microwave apparatus needed for ordinary standards is avoided, the drive signal being generated by the superconductive circuit itself. The only requirement for RSFQ circuits is an external accurate frequency reference, typically operating in the MHz range. The simplification in instrumentation must be traded off with a much higher complexity of superconductive circuit, namely an increased number of junctions, additional elements like inductors, and fabrication of junctions with different parameters in the same device. The basic element in RSFQ is the Josephson transmission line, a string of Josephson junctions connected by inductors where the pulse can propagate, like in transmission lines, and even amplified [21]. A basic Digital to Analog converter suitable for voltage standard will include at least some voltage multiplier stages, to increase the output signal to practical values. Coupling between Josephson transmission lines and voltage multipliers is obtained by capacitive coupling [22], but magnetic coupling through transformer-like circuits has proven to be more effective [23]. Nowadays DAC for Metrology are fairly more complex devices, with many digital blocks performing various specialized functions, and correspondingly high power consumption. Successful operation, of a 10-bit RSFQ DAC capable of generating up to 20 mV has recently been reported [24].

#### **4. Current junction technologies and fabrication issues**

The fabrication of arrays for AC generation and waveform synthesis is a difficult, challenging task, and many different technologies have been proposed and tested, yet the choice of material and fabrication techniques still represents an open question. Independent



of the technique adopted for signal generation, a relevant problem is posed by the high number of junctions needed to reach practical voltages, since junctions for AC typically operate on the first step and the drive frequency is limited for technical and economical reasons [25]. A high noise immunity, low power dissipation, reduced dimensions are also essential properties in metrological applications and in view of spreading quantum standards to a wider range of users.

The chip dimensions are set by the area and the number of junctions, both essential parameters for technology, because a reduced area along with a high number of junctions make it difficult to guarantee the uniformity of their electrical properties, which is essential to observe decent collective steps.

These requirements can be translated into well-defined specifications on junction parameters.

First, the critical current  $I_c$ , which sets an upper limit on the amplitude of the quantized steps, should be large enough for steps with suitable width and noise immunity, yet a too high value increases dissipation in junctions. The area of the junction must be small, to reduce array dimensions, but avoiding the excessive difficulties in fabrication imposed by deep submicron patterning, today still hard to achieve for high integration superconductive circuits. It follows that an optimal range exist for critical current density. Presently, values ranging from few kA to tens of thousands kA/cm<sup>2</sup> appear to be the best choice.

The characteristic voltage,  $V_c$  determines the microwave optimal drive frequency  $f_d$ , from the relation  $\Omega = f_d / f_c$ , where  $f_c$  is  $V_c \times (2e/h)$  and thus the step voltage (i.e. the voltage resolution of the array) and the number of junctions needed to achieve the maximum requested voltage output. To obtain the maximum step amplitude,  $\Omega$  must be  $\approx 1$  [26]. In order to use commercial microwave instrumentation and reduce as much as possible the number of junctions, achieving the maximum possible voltage output, drive frequencies close to 70 GHz are used, and  $V_c$  around 150  $\mu$ V are needed. Of course even larger values, which on the other hand are absolutely advantageous for speed applications, can be used. But in voltage standard application this causes a sensible reduction of the step width.

Moreover, the characteristic voltage defines also the highest speed of RSFQ circuits, since this is proportional to  $\tau_{car} = 1/f_c$ . In this case, the highest  $V_c$ , the higher the speed.

On the contrary, where we focus on the best possible voltage resolution of the standard, junctions with reduced  $V_c$  should be used.

In this chapter we discuss extensively niobium and niobium nitride based junctions, considering high  $T_c$  junctions in the section devoted to the use at temperatures above 4.2 K. A challenging problem to be solved in the next future for voltage metrology and superconductive electronics applications, is operation in cryocoolers at a temperature greater than 4.2 K. Indeed, in order to make the Josephson quantum standards available to a widespread market, rather than limited only to the National Laboratories, as well as affordable for private companies needing an accurate voltage reference, refrigeration systems cheaper and more compact than those nowadays used for niobium junctions are required. Moreover, as it will be discussed in a next section, the use of a cryocooler would allow a reduction of the measuring leads between the Josephson device and the measuring system, reducing the indetermination associated to the transition between voltage levels.

Considering the electrode configuration, non-hysteretic IV characteristic can be obtained by three main classes of junctions:

- hysteretic SIS with external shunt
- metallic barrier SNS
- double barrier SINIS.

The first class, directly derived from the most developed and optimized process of superconductive electronics, namely Nb/AlO<sub>x</sub>/Nb junctions and which is still the predominant technology for RSFQ circuits, suffers from the disadvantage of a configuration requiring an external resistor or a more complex circuitry. This, with the severe limitation to  $I_c$  from chaotic instabilities [27] limits the use of these junctions in voltage standard circuits. In spite of this a programmable array based on shunted Nb based SIS, where a reactive shunting was realized for blocks of several junctions, proved suitable operation at 1 V level [28].

In any case we will not consider junctions of this type in the following, where a detailed analysis limited to the last two classes will be carried out.

#### 4.1 Programmable voltage standards

Programmable voltage standards (PJAVS) are made with series-connected subarrays, whose size follows a power of two sequence. Through control of the bias current of each section, the series voltage is set by the code represented by on/off status of array bias lines. To switch subarrays, junctions that can generate a current-controlled, univocally defined voltage, i.e. junctions with a single valued (non-hysteretic) IV curve must be used. To provide signal to noise ratios adequate to metrological applications, critical currents at the mA level are at least required. On the other side, due to the fact that only the first Shapiro step of junctions is used, to achieve voltages of 1 and 10 V,  $V_c$  even in excess of 100  $\mu$ V are sought to reduce number of junctions.

As known, the amplitude of the  $n = 1$  quantized Shapiro step normalized to the critical current attains its maximum when the microwave drive frequency and the junction characteristic frequency are nearly equal. This corresponds also to a condition of minimal power dissipation [26]. This leads to optimal  $V_c$  of 100-200  $\mu$ V.

SINIS junctions achieve the shunting of the capacitance of a SIS junction, by using two, very thin oxide layers, separated by a metallic barrier. They are typically made of Nb/Al/AlO<sub>x</sub>/Al/AlO<sub>x</sub>/Al/Nb and have been extensively studied both theoretically and experimentally [29-34]. The best results on these junctions for superconductive electronics applications and especially for programmable voltage standard have been obtained by the PTB group, who realized complete successful circuits for programmable potentiometer at 1 and 10 V level with 7000 and 70.000 junctions respectively. These SINIS junctions feature  $J_c \leq 1$  kA/cm,  $V_c$  as high as more than 250  $\mu$ V at 4.2 K with junction size 100-1000  $\mu$ m<sup>2</sup>. SINIS junctions with various values of Al thickness and AlO<sub>x</sub> barrier transparency have been reported, but, even if they feature  $J_c$  and  $V_c$  higher than those given above, their IV characteristic was rather anomalous showing an high residual hysteresis at 4.2 K, and were hardly reproducible.

In general a problem for the SINIS junctions is due to the high transparency of the two insulating barriers, with the necessity to realize very smooth planarization of the underlying films, since the probability of defects such as pinholes is increased. The fabrication process, requiring that the two barriers and the three thin aluminum layers been highly homogenous, has set a limitation to the fabrication of really high number of junctions, as in arrays for programmable voltages at 10 V and more [35].

In SNS junctions the damping of the IV characteristic for non-hysteretic behavior is obtained by using a normal metal as barrier. SNS have high values of  $J_c$ , but typically the metals used have very low resistivities. Nb/PdAu/Nb junctions have been the first developed type of SNS junctions for this application by NIST group, and have produced stable voltage outputs at 1 V as binary arrays for programmable potentiometers, being also used in circuits for pulse driven AC synthesis [14],[15]. They feature values in the range of 100 kA/cm<sup>2</sup> and  $V_c$  between 5 and 30  $\mu$ V with typical size of few  $\mu$ m<sup>2</sup>, even if  $J_c$  values would support also submicron dimensions. The values of  $V_c$  limit their use to drive frequencies of few GHz and circuits have about 30.000 junctions for 1 V.

Impressive results have been obtained with NbN/TiN/NbN junctions, developed by the Tsukuba group, since very large arrays, with more than 300.000 series connected junctions have been successfully tested in a 8 and 11 bit DAC circuits with quantized steps at 10 V at 10 K [36],[37]. Features of these junctions are  $J_c$  about 10 to 10 A/cm and  $V_c$  10-20  $\mu$ V at 10 K with areas of few square micrometer. 20 V output has been also achieved, by series connecting two arrays of different chips, minimizing interconnecting dissipation [9].

Materials have been also proposed as normal metal barrier which are at the metallic insulator transition. In such a way it was possible to tune the barrier resistivity and the characteristic voltage of the junctions [38],[39]. Among the experimental results we mention Nb/TaO<sub>x</sub>/Nb, NbN/TaN/NbN and Nb/Nb<sub>x</sub>Si/Nb junctions. Nb/TaO<sub>x</sub>/Nb junctions fabricated at IEN-INRiM, showed resistivities varying different order of magnitudes depending on the bias voltage of the cathode during the sputtering deposition of the TaO<sub>x</sub> film [40]. The junctions have however a marked aging, featuring a reduction of  $J_c$  at values less than 10 A/cm. Also NbN/TaN/NbN junctions studied in [41] have not yet been suitable for large circuits production. These junctions featured very high  $V_c$ , up to more than 0.5 mV at 4.2 K with  $J_c$  10<sup>4</sup>-10<sup>5</sup> A/cm<sup>2</sup>, but required a difficult tuning of fabrication parameters and also the temperature dependence of these parameters was critical.

In this category, Nb/Nb<sub>x</sub>Si<sub>1-x</sub>/Nb junctions represent the most successful attempt so far. Although previously studied by Barrera and Beasley in the '80 [42], the most promising results have been achieved in the last years by the NIST researchers, which experimented these junctions in arrays at 1 V level [43]. These authors achieved a transition from a conductive to an highly resistive phase by varying the sputter deposition power of the two elements.  $J_c$  varying from 10 to 10 kA/cm<sup>2</sup> and  $V_c$  from few  $\mu$ V to 150  $\mu$ V and more at 4.2 K have been obtained. The stabilization of the barrier stoichiometry, critical aspect of this type of junctions, can be achieved by a thermal annealing after the deposition of the trilayer. Depending on the Nb content and the barrier thickness, also the temperature stability of these junctions can be optimized, see in a next section [44].

## 4.2 Pulsed standards for waveform synthesis

The series of pulses used for waveform synthesis has a very rich harmonic content and requires proper optimization of the microwave behavior in circuits for rf signal distribution. Most of the techniques adopted for analysis and design of circuits with distributed parameters assumes operation with narrow band signals, thus cannot be directly applied for pulse transmission. The most reliable solution to guarantee pulse operation is to reduce array dimensions allowing lumped analysis of circuit. To that aim array length must be shorter than  $\lambda/8$ , where  $\lambda$  is the wavelength of the microwave signal. This constraint sets strong requirements on junctions technology, in particular on junction size.

Therefore the main stream for the realization of these devices is a) submicrometric junction technology development or/and b) vertical multistacking of junctions

In order to realize the array configuration described, materials resulting in improved steepness of the etch profile are preferred, especially when multistacks of junctions are employed.

SINIS junctions, initially developed for programmable voltage standards and previously described, have also been tried for AC synthesis, notwithstanding their limitations [33].

However, as mentioned above, the first, and presently the best results in pulse driven standards had been attained with Nb/PdAu/Nb junctions, which, due to their high current density allow to strongly reduce their dimensions [14].

Nb/HfTi/Nb junctions have been experimented by PTB team with successful results in the realization of large submicrometric circuits for AC synthesis which have recently reached quantized voltages from 2000 series junctions. These junctions feature about 80 kA/cm for HfTi 20 nm thick and of 15  $\mu\text{V}$  in junctions  $0.2 \mu\text{m} \times 0.2 \mu\text{m}$  [33].

Other type of SNS junctions with similar low characteristic voltage, such as Nb/Ti/Nb junctions developed by Jena team and Nb/Al/Nb junctions studied at INRiM have not been resulted in large circuits yet [45]. Ti barrier junctions feature of  $10^4$ - $10^5$  A/cm<sup>2</sup> and from 10  $\mu\text{V}$  up to 100  $\mu\text{V}$   $V_c$ , while for Al film 100 nm thick, also higher values of and similar or lower values for were obtained. In both case the junction size was in the micrometer and submicrometer range.

Nb/MoSi/Nb junctions also studied and experimented by NIST team, have been used to realize multistacked junctions for lumped arrays for AC synthesis [46]. They feature critical current density between  $10^4$  and  $10^5$  A/cm<sup>2</sup>, depending on barrier thickness, while characteristic voltage is typically few tens of  $\mu\text{V}$  in junctions of few  $\mu\text{m}^2$ .

## 4.3 Power control standards

Power control standards set some peculiar requisites on junction technologies, their operation being based on the overlapping of the range of currents of the steps and critical current. To fulfill these requirements, a tight control of junction hysteresis is needed.

Referring in particular to [47], where the parameters which must be fulfilled by the junctions to obtain the overlapping of the  $n=0$  and  $n=1$  step are discussed, it is pointed that the mentioned overlapping is obtained in a stable way when  $f > (1.5-2) f_c$  [48]. In particular the authors suggest the employ of edge type junctions with  $V_c$  as high as 1 mV.

However, in principle, no constraint on  $V_c$  is needed, while the need of a sufficient, but not exceedingly high overlapping of the two steps is the only fundamental aspect.

#### 4.4 Application to RSFQ

The application of superconductive technology to RSFQ requires a high level of reliability, suitable for the development and testing of very complex circuits, with thousands of junctions implementing many different functional blocks. The best established technology, well mastered by many foundries, is based on SIS junctions, usually made of a Nb/AIOx/Nb, with an externally shunt resistor [49]. The minimum junction area that can be fabricated with this technology is  $12 \text{ } \mu\text{m}^2$ , and the whole process involves as many as 12 mask steps. Impressive results have been obtained with the standard Nb/AIOx/Nb technology: in [50] the operation of a RSFQ quantum DAC with 6000 Josephson junctions, subdivided in several functional blocks, is reported. SINIS junctions have proven to be suitable for RSFQ, and some fundamental circuits based on this technology were successfully operated [51]. More recently, owing to the wide range of tunability of their electrical parameters, co-sputtered Niobium-silicide barrier junctions appear interesting for a wide range of applications in digital electronics and well suited to RSFQ [52]. Even more complex circuits, comprehending RSFQ electronics as a pattern generator of pulses, followed by semiconductor amplifier and a Josephson junction series array as a quantizer are studied, and realization of preliminary building blocks is in progress.

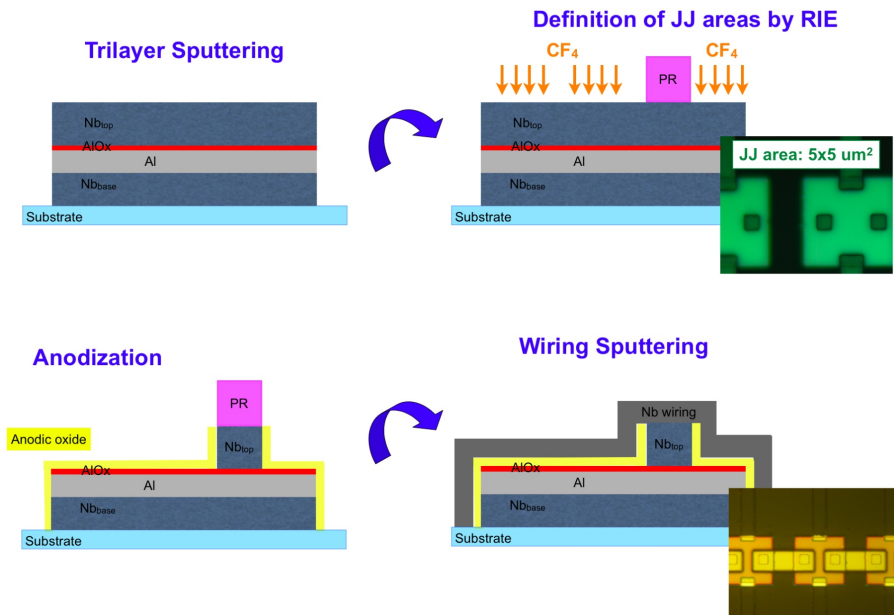


Fig. 6. Example of fabrication process sequence of a Josephson junctions array. The process is started with the trilayer deposition, by sputtering, followed by the area patterning, by photolithographic process and reactive ion etching, and the insulation of the electrodes by liquid anodization. The deposition of contact layers concludes the process.

#### 4.5 An example of a multifunctional junction: Nb/Al-AIOx/Nb overdamped SNIS junctions

SNIS Josephson devices of the Nb/Al-AIOx/Nb type do belong to the family of internally shunted junctions, exhibiting important similarities, but demonstrating additional advantages comparing with other types of self-shunted junctions, namely an extended range of electrical parameters and an advanced temperature stability of transport characteristics.

The main difference respect to the structure of the basic Nb/AlOx/Nb SIS junctions is the thickness of the aluminum film which is increased at tens of nm, up to about 100 nm, and the augmented transparency of the AlOx [53],[54].

An essential feature of SNIS is that, at 4.2 K a transition from the hysteretic to the non-hysteretic state can be induced, when the aluminum thickness is in the above mentioned range, by changing the AlOx exposure dose. Conversely, once chosen Al thickness and AlOx exposure values, this transition is observed as function of the junction temperature in measurements below 4.2 K (see Figure 7) [63].

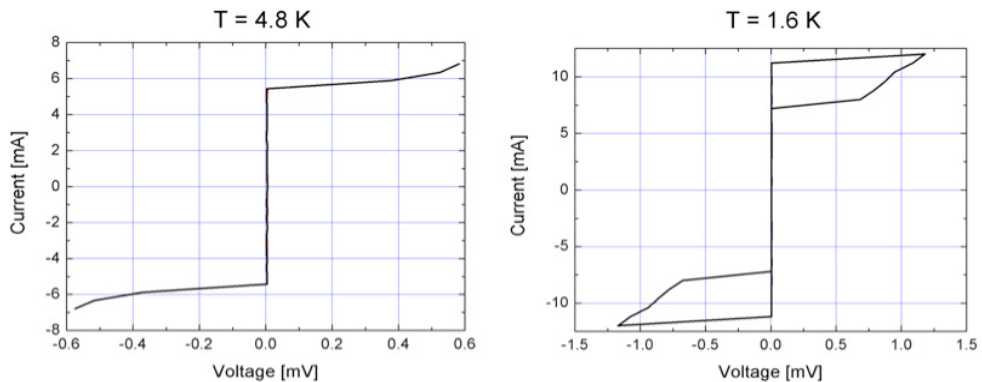


Fig. 7. Transition from non-hysteretic to hysteretic behavior for a SNIS junction, obtained by changing the operating temperature.

Critical current densities  $J_c$  of the SNIS junctions are sufficiently high (also  $2 \text{ mA}/\mu\text{m}^2$ , along with characteristic voltage values  $V_c$  up to several hundred microvolts at 4.2 K. In particular, junctions with Al-layer thickness ranged from 40 to 120 nm and exposure doses of AlOx between 160 and 250 Pa·s showed  $J_c$  from 0.01 to  $1 \text{ mA}/\mu\text{m}^2$  and  $V_c$  from 0.1 to 0.7 mV at 4.2 K. The dependence of  $J_c$  on  $E_{\text{ox}}$  within the indicated interval of  $E_{\text{ox}}$  is almost linear in the log-log scale, as for other types of Josephson devices with oxide barriers, while the characteristic voltages are more dependent on  $d_{\text{Al}}$ .

The specific normal conductance of these devices ranges from  $0.3 \cdot 10^8 \text{ Ohm}^{-1}\text{cm}^{-2}$  to  $0.7 \cdot 10^8 \text{ Ohm}^{-1}\text{cm}^{-2}$  and, hence, is of the same order of magnitude as that in high-current-density single-barrier Josephson junctions and double-barrier SINIS heterostructures [54].

As indicated in the literature, the main explanation of the self-shunting phenomenon in high specific-conductance devices both SIS and in SINIS structures, consists in the enhanced subgap conductivity. The same result has been verified in SNIS junctions where the subgap

resistance  $R_{sg}$  at 4.2 K and very low voltages is of the order of 0.3 Ohm which is higher but comparable with the normal-state junction resistance  $R_N$ .

The main contribution to this subgap enhancement may be related to Andreev reflection processes in the interlayer. Really, since the Al film thickness in SNIS junctions is comparable with the electron mean free path, the transparency of the weak link between the Nb electrodes is dominated by that of the insulating layer. Its average resistance nearly coincides with that in high-specific-conductance Nb-AlO<sub>x</sub>-Nb junctions where appearance of multiple Andreev reflections was revealed by analyzing comparatively high subgap 'leakage' currents in Nb-AlO<sub>x</sub>-Nb junctions. Hence, the explanation of the phenomenon suggested for these structures and based on the existence of a universal distribution of transparencies in dirty disordered tunnel barriers [55],[56] in the interlayer. Really, since the Al film thickness in SNIS junctions is comparable with the electron mean free path, the transparency of the weak link between the Nb electrodes is dominated by that of the insulating layer, Its average resistance nearly coincides with that in high-specific-conductance Nb-AlO<sub>x</sub>-Nb junctions where appearance of multiple Andreev reflections was revealed by analyzing comparatively high subgap 'leakage' currents in Nb-AlO<sub>x</sub>-Nb junctions. Hence, the explanation of the phenomenon suggested for these structures and based on the existence of a universal distribution of transparencies in dirty disordered tunnel barriers is appropriate to these devices as well. They do not arise from trivial pinholes because after suppressing the Josephson effect, usually no leakage supercurrents were observed.

## 5. Special issues for next generation standard and possible solutions

### 5.1 About the extended use of Josephson Voltage standard in AC measurement

The use of Josephson junctions arrays for AC measurement, while has already set up reference standards of AC voltages in defined ranges of amplitude and frequency, as described in section 2, represents also the first step of new quantum based electrical metrology and instrumentation, of which the quantum voltmeters represent an example.

A first application to AC measurement has been realized by using a programmable array to generate the AC voltage to be compared, at some frequencies, with the DC voltage of a thermal converter, through a suitable balance bridge. In such a way the determination of this standard is also linked to the  $e$  and  $h$  fundamental constant. A comparison between the traditional technique and the one employing PJVS has demonstrated and the accuracy for AC voltages up to 100 Hz can be better than one part in  $10^7$ .

Further improvement are expected concerning AC equivalent of the resistance: it is in fact possible to realize quantum impedance bridges, by using two separate sections of an array, or even two different arrays each separately radiated by a microwave signal, with a known and measurable phase difference. A two terminal pair impedance bridge has achieved an accuracy of one part in  $10^8$  in a frequency range from 25 Hz to 10 KHz, limited mainly by contact resistance changes at the bridge terminals [57]. Further progress is expected when a four terminal connection to the resistances can be provided..

Also quantum sampling of waveforms is made possible, by realization of Josephson DAC with quantum accuracy, and using them with a proper electronics which enclose in the same circuit conventional and quantum based DAC.

A further employ, presently already in progress, is the realization of standards for Johnson noise thermometry, where arrays of a small number of Josephson junctions are used to generate random white noise voltages. For this application, while the amplitude is not important, the device should extend the frequency range of quantum standard as much as possible [58].

All these aspects have been the subject of specific research projects of Euromet, named ProVolt, JAWS, Binary Josephson Array Power Standard, JOSY and future proposals of next calls.

But, considering the use of Josephson junctions for Voltage standard, the content of this chapter, what are the main goals for the next generation devices?

As mentioned above, see figure 7, both amplitude and frequency of the AC voltage must be pushed to higher values.

It should of course be observed that for people working in the AC field the value of importance is the rms voltage, while typically the output from the values mentioned for some of the mentioned devices is the maximum output DC value.

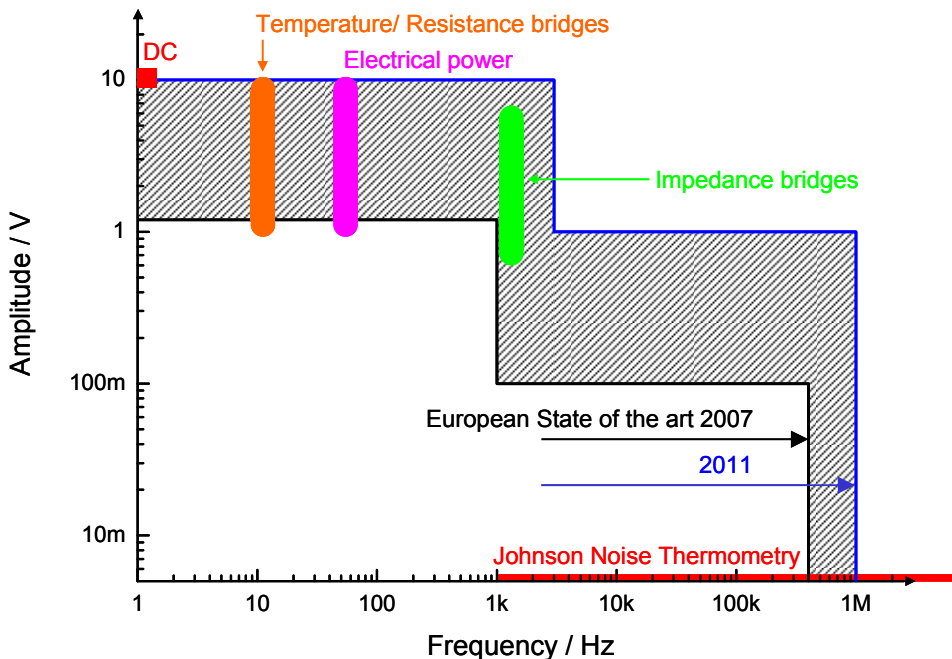


Fig. 8. Josephson programmable and pulse driven AC Voltage standard impact on measurement of AC voltages for some applications. The amplitude and the frequency extension for the time period from 2007 to 2011 is indicated. (from IMERA TP 4 Josy project extended summary).



To increase the output voltage of the arrays is a serious challenge since nowadays to have more than  $10^6$  Josephson junctions on the same chip with tolerable spread of parameters (5-10%), is beyond the state of art of superconductive technology, As reported in a previous section, 250.000 junctions in plane or 330.000 in vertical stacks of three junctions [17], are the best result sofar.

So, waiting for a technology which can overcome this limit a different approach must be presently searched, reducing as much as possible the number  $N$  of junctions and their dimensions for a given voltage output, while keeping the desired voltage resolution.

Another relevant issue is in the possibility of operation at temperatures higher than 4.2 K, in view of the substitution of expensive liquid helium refrigeration systems with the compact cryocoolers which can lead the diffusion of voltage standards to the private companies.

This is a challenge of great importance for superconductive circuits and voltage standards apparently not to be solved in a short period by high  $T_c$  junctions. At present, large arrays fabricated with higher critical temperature superconductors, like YBCO or the more recent  $MgB_2$ , are not yet available, since the technology of these junctions do not allow to achieve the required integration level [59]. One big problem of these junctions is the stability in time and with thermal cycling.

The most interesting results sofar have been achieved with YBCO bicrystal shunted junctions, where quantized steps have been measured near 77 K above 100 mV [60].

Concerning niobium and niobium nitride based junctions, only NbN/TiN/NbN have demonstrated operation at these temperatures, achieving a sound result such as a 11 bit DAC with 10 V output at 10 K, which were also risen at 20 V, by using two separate chip connected in series [9].

However, they are sensitive to temperature changes, requiring a stabilization of the cryocooler at 0.1 K level, have a strong demand on dissipated power and require a top-level, costly, fabrication process.

The need of the temperature stability means that small temperature variations should cause only small current changes and can be of interest for applications where a simplified refrigerator is used such as RSFQ and also voltage standard applications [19]. For example, conventional SIS junctions exhibit an excellently stable temperature range from 0 to nearly  $0.6 T_c$  but a very strong  $I_c$ -vs- $T$  dependence above.

A possible solution for both the above mentioned issues is to use high characteristic voltage junctions.

As reported in section 3, this is not trivial, since intrinsically overdamped junctions have  $V_c$  below 100  $\mu$ V at 4.2 K.

However two different technologies have been recently proposed able to produce  $V_c$  as high as 0.5 mV and more at such temperature: Nb/Nb<sub>x</sub>Si<sub>1-x</sub>/Nb SNS and Nb/Al-AlO<sub>x</sub>/Nb SNIS junctions

In the following results using SNIS are reported.

## 5.2 Advantages of using junctions with high characteristic voltage: a) Use of higher order steps to optimize $V_{\max}/N$

A specific feature of SNIS junctions is the possibility of achieving high values of both  $I_c$  and  $V_c$ , see section 3.5, with current density  $J_c$  up to  $0.5 \text{ mA}/\mu\text{m}^2$  and  $V_c$  up to  $0.7 \text{ mV}$  at  $4.2 \text{ K}$ .

It is then possible to fabricate junctions with dimensions in the range of  $1 \mu\text{m}^2$ , with critical currents in the mA range, which ensures that adequate amplitudes of the Shapiro steps are obtained in circuits with reduced dimensions.

Concerning the step amplitude, as reported in section 3.1, the optimal value for the Shapiro steps in overdamped junctions is obtained when  $f_{\text{drive}} \approx f_c = 2e/h V_c$ , which gives values of about  $0.15 \text{ mV}$  for  $f_{\text{drive}} = 70 \text{ GHz}$ . The choice of  $f_{\text{drive}} = f_c$  is regarded as the best operating condition for Josephson binary arrays since it ensures the minimal demand of microwave power for equal and maximized 0 and  $\pm 1$  step amplitude. In fact for many applications the step  $n = 0$  is needed to establish zero output voltage. However, steps corresponding to higher order harmonics are enhanced for  $f_c$  multiple of  $f_{\text{drive}}$  [61]. This allows to measure voltages twice, three and also four times higher for a given number of junctions and for instance, a binary-divided arrays consisting of 8192 SNIS Josephson junction has given output voltages of 1.25, 2.5, 3.75 and  $5 \text{ V}$  [62].

Of course an increased power dissipation is required to work on higher order steps, proportional to the square of the bias current of the step, especially if the maximum amplitude of the step is searched.

In figure 9 the amplitude of the Bessel functions of order 1 to 3, to which are proportional to the corresponding steps, is plotted vs. the argument of these function, which, according to Shapiro calculations, is related to the voltage of the microwave signal of frequency  $f_{\text{drive}}$

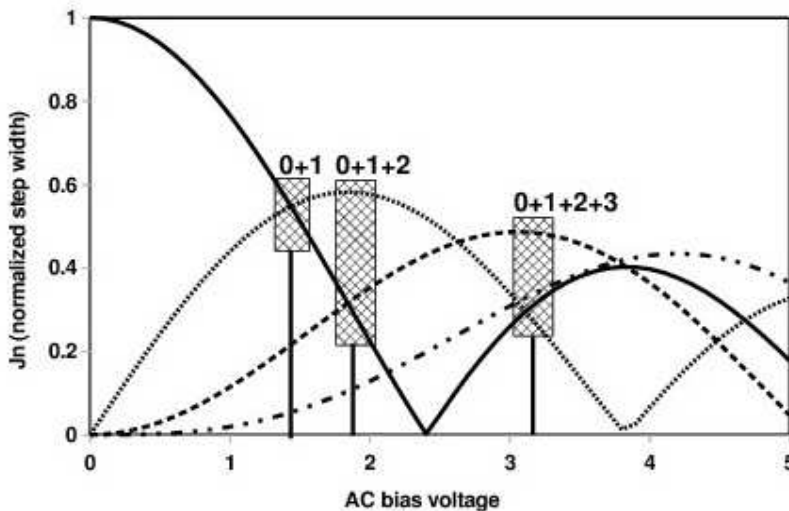


Fig. 9. Step amplitude behavior for  $n=0-3$ , see eq. 4. Conditions for optimal 0 and 1, 0, 1 and 2 and also 0, 1, 2 and 3 are indicated.

And it must be remembered that, when higher order steps are used, the voltage resolution of the device is correspondingly reduced, since now  $v_{\min.} = f_{\text{drive}} \times n$ .

Therefore it must be evaluated what is the target to be fulfilled: if maximum voltage output, circuit dimensions, power dissipation or resolution. In correspondence different type of standard can be fabricated.

An optimal solution seems achieved when using arrays made of junctions with characteristic voltage in the range between  $f_{\text{drive}} < f_c < 2f_{\text{drive}}$ , where it is possible to have both first and second step enhanced.

From figure 9 the condition where the steps  $n=0, 1$  and  $2$  are all optimized requires a low increment of the microwave voltage, with respect to that for optimized steps  $n=0$  and  $n=1$ .

Experimentally, for a radiation frequency of 70 GHz, an RF power of about 40 mW, measured at the input flange of the waveguide cryoprobe for equal 1st and 2nd step (cf. Figure 10), is slightly higher of 20-25 mW for the case of  $n=0$  and  $n=1$ . The  $V_c$  of this array was 0.3 mV. The step amplitudes, measured with oscilloscope and with sub-microvolt techniques confirm flatness at metrological level (Figure 9 inset), while the RF power is slightly higher of 20-25 mW for the case of  $n=0$  and  $n=1$ . The  $V_c$  of this array was 0.3 mV. The step amplitudes, measured with oscilloscope and with sub-microvolt techniques confirm flatness at metrological level (Figure 10).

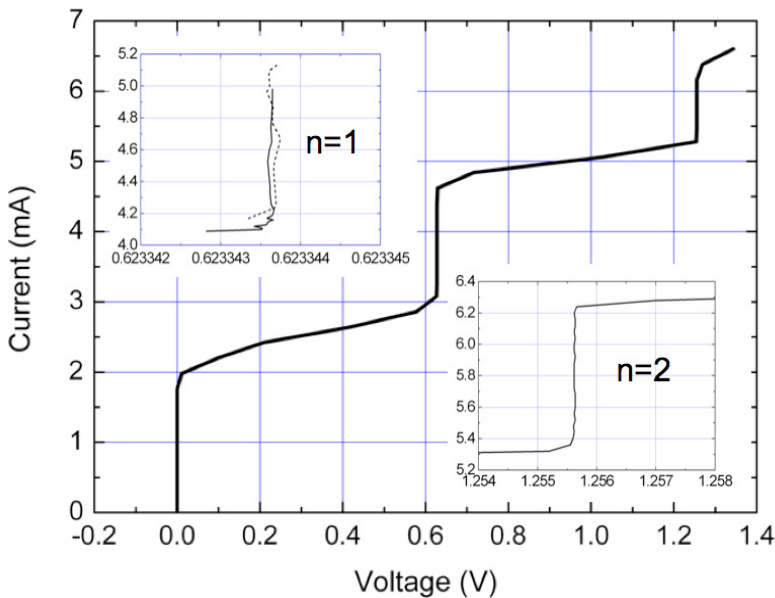


Fig. 10. Part of a binary-divided array consisting of 4096 SNIS junctions irradiated by 73 GHz microwaves. The microwave power, 40 mW at the input flange, has been optimized to have equal and wide (1 mA) 1st and 2nd steps. The insets show high resolution measurements of steps profiles. The  $n=1$  step was traced with a precision DVM (dashed) and by array comparison (continuous).

The precision measurement have been carried out using different approach: i) the array voltage has been measured with a Digital Voltage Multimeter, DVM, using multi averaging to reduce noise, ii) the voltage of one half of the array was measured vs. the voltage of the other half, biased on the same step, iii) the voltage of the programmable array was compared with the 10 V DC voltage standard, biased at the appropriate voltage level. Depending on the method, the resolution ranged between 1 $\mu$ V and 10 nV.

Being  $V = Nn (h/2e) f_{drive}$ , where  $N$  is the number of series connected junctions and  $n$  is the step order, this results in the possibility of using an array where half the number of junctions provide the same voltage output, with the same resolution, by realizing a suitable biasing electronics for 1<sup>st</sup>, and 2<sup>nd</sup> and 0 step.

As a consequence, the chip size can be reduced of a factor 2 respect to the size of the present type of programmable voltage standard .

### 5.3 Advantages of high characteristic voltage: b) Operation of voltage references above 4.2 K

The high value of  $V_c$  at 4.2 K is also a tool for operation of the arrays above 4.2 K, since the scaling of the electrical parameters with temperature, still allows to use them at higher temperatures.

However a further requirement to the junctions is set by the temperature stability of the electrical parameters, especially above 4.2 K. For a superconductive junctions in fact several non-equilibrium phenomena arise when the operating temperature approaches  $T_c$

The stability of  $V_c$  mainly depends from the critical current dependence on temperature, since the junction resistance is almost constant in that temperature range for the majority of metallic and insulating barriers. In particular an evaluation of the temperature stability can be made considering the temperature derivative of the current, normalized to the value at  $T = 0$  , vs. the incremental temperature, normalized to  $T_c$ :  $d(I_c(T) / I_c(0)) / d(T / T_c)$

In figure 11 the relation is plotted for different type of Josephson junctions.

For hysteretic SIS junctions, whose temperature behavior is described according to the Ambegaokar and Baratoff theory, curve AB, a strong change is observed especially above 4.2. On the contrary for SNS junction, described by a Kulik and Omelianchuk formulation, curve KO, a less pronounced dependence can be observed, but it must not be ignored that in the majority of case these junctions have quite low  $V_c$  values already at 4.2 K [63].

In the figure are also shown three calculated curves for SNIS junctions, where by changing the N/S relative thickness in the Nb/Al first electrode, through the parameter of the parameter  $\gamma_{eff}$ , the temperature dependence can be properly tailored.

In particular, the dotted curve, corresponding to a thickness of the aluminum film about 100 nm, shows a a very small change from 0.7 T/ $T_c$ .

Thus, it is possible to engineer the temperature stability of 4-layered structure by increasing the thickness of the different layers and changing, as a result, the shape of supercurrent-vs-temperature characteristics [64].

Experiments on single junctions and small array of 10 junctions demonstrated the operation up to temperatures near junction  $T_c$  [65].

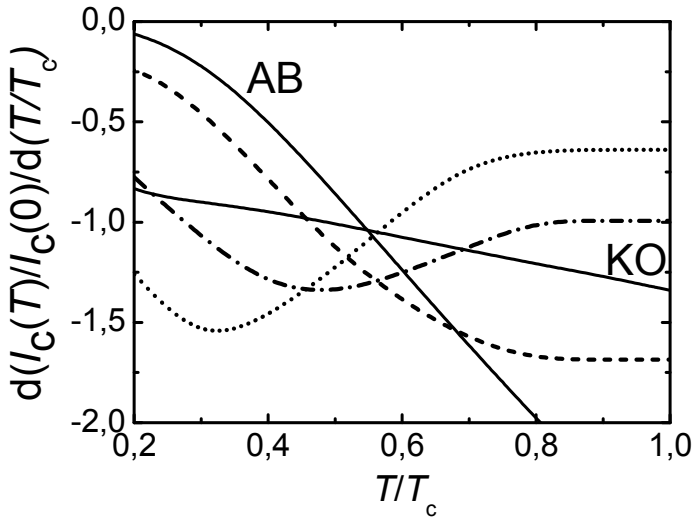


Fig. 11. Normalized temperature derivative of the critical supercurrent calculated for conventional SIS (the AB curve) and SNS (the curve KO) junctions and the SNIS devices with  $\gamma_{\text{eff}} = 2, 10,$  and  $20$  (dashed, dashed-dotted, and dotted lines respectively).

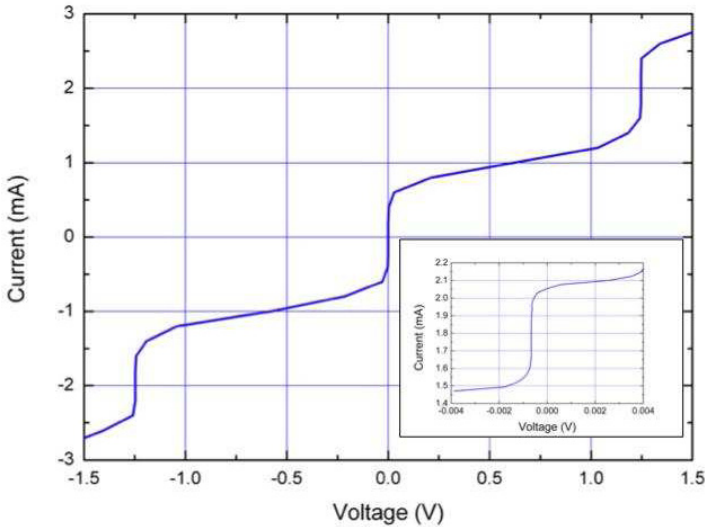


Fig. 12. Voltage step at 1.25 V measured at  $T = 6.3$  K on a binary-divided array of 8192 JJ. The inset shows detail of the quantized step, 0.25 mA wide.

Of course when using more complex circuits, as the programmable arrays described above, non-uniformities in the fabrication and in the distribution of the microwave signal reduce the maximum temperature, especially if the device is operated in gas.

Figure 12 show steps at 1.25 V, 0.25 mA of a binary array, measured at  $T = 6.3$  K Wider steps, 0.5 to 1 mA, have been measured up to 7.2 K on subsections of the arrays.

It is expected that the results will be improved by improving fabrication homogeneity, microwave signal coupling and using a cryocooler setup.

## 6. Conclusions

The extension of DC Josephson Voltage standard to AC has led to the development of different type of circuits, according to the specific application, which have been analyzed in section 2.

In section 3 a comparison of different types of overdamped junctions developed for the large circuits needed by the AC extension of Josephson voltage standard has shown advantages and disadvantages of a certain electrode configuration or material. In order to achieve the desired noise immunity of the quantized voltage levels, the junctions must maximize both  $J_c$  and  $V_c$  this last issue being also useful for applications where the highest speed of the Josephson circuits is expected. And these issues must be attained using a tolerable degree of technological effort, also in terms of allowable integration. In fact, in order to have a widespread development in the world laboratories for these circuits, neither the number of junctions should be increased nor the size reduced beyond a certain limit. All together, these issues are not solved by any of the junctions reported in this analysis, even if some of the described technologies had achieved sound results.

In section 4 future developments of these circuits have been considered with attention to the problems to be solved.

Waiting for a further development of superconductive junction technology integration, and the use of high temperature materials, the employ of high characteristic voltage junctions is presented as a possible route for extending the performances and applications of Josephson voltage standard not only in fundamental metrology but also for laboratory instrumentation.

## 7. References

- [1] J. Bardeen, L.N. Cooper, J.R. Schrieffer, Theory of Superconductivity, *Phys. Rev.*, 108(5): 1175-1204, 1957S. Tsai, A.K. Jain, J.E. Lukens, High-precision test of the universality of the Josephson voltage-frequency relation, *Phys. Rev. Lett.*, 51(4):316-319, 1983.
- [2] B.D Josephson, Possible new effects in superconductive tunnelling, *Physics Letters*, 1(7): 251-253, 1962.
- [3] A.F. Andreev, Thermal conductivity of the intermediate state of superconductors, *Sov. Phys. JETP* 19: 1228, 1964.
- [4] M. T Levinsen, R.Y. Chiao, M.J. Feldman, B.A. Tucker, An inverse ac Josephson effect voltage standard, *APL*, 31, 11, 776 (1977).
- [5] C.A Hamilton, C.J. Burroughs, and R.L Kautz. Josephson d/a converter with fundamental accuracy. *IEEE Trans. Instr. Meas.*, 44(2):223-225, 1995.
- [6] R. Behr, J. Kohlmann, J.T.B.M. Janssen, P. Kleinschmidt, JM Williams, S. Djordjevic, J.P. Lo-Hive, F. Piquemal, P.O. Hetland, D. Reymann, et al. Analysis of different measurement setups for a programmable Josephson voltage standard. *IEEE Trans. Instr. Meas.*, 52(2):524-528, 2003.
- [7] G. Geneves, P. Gournay, A. Gosset, M. Lecollinet, F. Villar, P. Pinot, P. Juncar, A. Clairon, A. Landragin, D. Holleville, et al. The BNM watt balance project. *IEEE Trans. Instr. Meas.*, 2005.

- [8] C.J. Burroughs, S.P. Benz, P.D. Dresselhaus, B.C. Waltrip, T.L. Nelson, Y. Chong, J.M. Williams, D. Henderson, P. Patel, L. Palafox, et al. Development of a 60 Hz Power Standard Using SNS Programmable Josephson Voltage Standards. *IEEE Trans. Instr. Meas.*, 56(2):289, 2007.
- [9] H. Yamamori, T. Yamada, H. Sasaki, and A. Shoji. A 10 V programmable Josephson voltage standard circuit with a maximum output voltage of 20 V. *Superconductor Science and Technology*, 21(10):105007, 2008.
- [10] J. Kohlmann, R. Behr, and T. Funck. Josephson voltage standards. *Measurement Science and Technology*, 14(8):1216–1228, 2003.
- [11] R. Monaco. Enhanced AC Josephson effect. *Journal of Applied Physics*, 68:679, 1990.
- [12] J-H. Kim A. Sosso, A.F. Clark Dynamics of overdamped Josephson junctions driven by a square-wave pulse, *JAP*, 83,6, 3225-3232 (1998).
- [13] J-H. Kim A. Sosso. Phase lock of non-hysteretic Josephson junctions with pulse bias: analytical properties, *Phil. Mag. Part B*, 1463-6417, 80, 5, 973 – 977 (2000).
- [14] S.P. Benz and C.A. Hamilton. A pulse-driven programmable Josephson voltage standard, *Applied Physics Letters*, 68:3171, 1996.
- [15] S.P. Benz, C.A. Hamilton, C.J. Burroughs, and T.E. Harvey. AC and DC bipolar voltage standard using quantized pulses. *IEEE Trans. Instrum. Meas.*, 48:266–269, 1999.
- [16] J. Kohlmann, O.F. Kieler, R. Iuzzolino, J. Lee, R. Behr, B. Egeling, and F. Muller. Development and investigation of SNS Josephson arrays for the Josephson arbitrary waveform synthesizer. *IEEE Trans. Instr. Meas.*, 58(4):797–802, April 2009.
- [17] S.P. Benz. Synthesizing accurate voltages with superconducting quantum-based standards, *IEEE Instrum. Magaz.*, vol.13, no.3, pp.8-13, June 2010.
- [18] A.S. Katkov, A.M. Klushin, G.P. Telitchenko, R. Behr, and J. Niemeyer. Challenges of Josephson junction arrays for AC voltage generation by microwave pulse power modulation. *IEEE Trans. Appl. Supercond.*, 15(2 Part 1):352–355, 2005.
- [19] H.J.M. Ter Brake, F.I. Buchholz, G. Burnell, T. Claeson, D. Cr  t  , P. Febvre, GJ Gerritsma, H. Hilgenkamp, R. Humphreys, Z. Ivanov, et al. SCENET roadmap for superconductor digital electronics. *Physica C: Superconductivity and its applications*, 439(1):1–41, 2006.
- [20] V.K. Semenov. Digital to analog conversion based on processing of the SFQ pulses. *IEEE Transactions on Applied Superconductivity*, 3(1 Part 4):2637–2640, 1993.
- [21] K.K. Likharev and V.K. Semenov. RSFQ logic/memory family: a new Josephson-junction technology for sub-Terahertz-clock-frequency digital systems. *Applied Superconductivity, IEEE Transactions on*, 1(1):3–28, Mar 1991.
- [22] V. K. Semenov and M. A. Voronova. DC voltage multipliers - A novel application of synchronization in Josephson junction arrays. *IEEE Transactions on Magnetics*, 25:1432–1435, March 1989.
- [23] C.A. Hamilton. Josephson voltage standard based on single-flux-quantum voltage multipliers. *IEEE Transactions on Applied Superconductivity*, 2(3):139–142, 1992.
- [24] M. Maezawa and F. Hirayama. 10-bit rapid single flux quantum digital-to-analog converter for AC voltage standard. In *Journal of Physics: Conference Series*, volume 97, page 012161. Institute of Physics Publishing, 2008.
- [25] J. Niemeyer. Josephson arrays for DC and AC metrology. *Superconductor Science & Technology*, 13(5):546–550, 2000.
- [26] R.L. Kautz, Shapiro steps in large-area metallic-barrier Josephson junctions, *J. Appl. Phys.*, 78, 9, (1995).
- [27] R.L. Kautz, R. Monaco, Survey of chaos in the rf-biased Josephson junction, *Journal of Applied Physics*, 57(3): 875-889, 1985.

- [28] J. Hassel, H. Seppa, L. Gronberg, I. Suni, SIS junctions with frequency dependent damping for a programmable Josephson voltage standard, *IEEE Trans. Instrum. Meas.*, 50(2):195-198, 2001.
- [29] M.Yu. Kupriyanov, A. Brinkman, A.A. Golubov, M. Siegel, H. Rogalla, *Physica C Superconductivity*, 326-327: 16-45, 1999
- [30] L. Capogna and M.G. Blamire. Superconducting proximity effect through high-quality high-conductance tunnel barriers. *Physical Review B*, 53(9):5683-5687, 1996.
- [31] F. Mueller, R. Behr, L. Palafox, J. Kohlmann, R. Wendisch, and I. Krasnopolin. Improved 10 V SINIS Series Arrays for Applications in AC Voltage Metrology. *IEEE Trans. Appl. Supercond.*, 17(2):649-652, 2007.
- [32] R. Behr, J. M. Williams, F. Patel, B. M. Janssen, T. Funck, and M. Klonz. Synthesis of precision AC waveform using a sinis Josephson array. *IEEE Trans. Instr. Meas.*, 54:612-615, 2005.
- [33] J. Kohlmann, F. Mueller, O. Kieler, R. Behr, J. Palafox, M. Kahmann, and J. Niemeyer. Josephson Series Arrays for Programmable 10-V SINIS Josephson Voltage Standards and for Josephson Arbitrary Waveform Synthesizers Based on SNS Junctions. *IEEE Trans. Instrum. Meas.*, 56(2):472-475, 2007.
- [34] F. Born, D. Cassel, K. Ilin, A. M. Klushin, M. Siegel, A. Brinkman, A. A. Golubov, M. Yu Kupriyanov, and H. Rogalla. Transport properties of SINIS junctions with high-current density. *IEEE Trans. Appl. Supercond.*, 13(2):1079-1084, 2003.
- [35] F. Mueller, R. Behr, T. Weimann, L. Palafox, D. Olaya, P.D. Dresselhaus, S.P. Benz, 1 V and 10 V SNS Programmable Voltage Standards for 70 GHz, *IEEE Transactions on Applied Superc.* 975 - 980, (2009). 10.1109/TASC.2009.2017911.
- [36] H. Yamamori, M. Ishizaki, A. Shoji, P. D. Dresselhaus, and S. P. Benz. 10 V programmable Josephson voltage standard circuits using NbN/TiN/NbN/TiN/NbN double-junction stacks. *Appl. Phys. Lett.*, 88(4):042503-1-042503-3, 2006.
- [37] H. Yamamori, M. Ishizaki, H. Sasaki, and A. Shoji. Operating Margins of a 10 V Programmable Josephson Voltage Standard Circuit Using NbN/TiN/NbN/TiN/NbN Double-Junction Stacks. *IEEE Trans. Appl. Supercond.*, 17(2 Part 1):858-863, 2007.
- [38] J. K. Freericks, B. K. Nikolic, and P. Miller. Tuning a Josephson junction through a quantum critical point. *Phys. Rev. B*, 64(5):054511-1-054511-13, 2001.
- [39] J.K. Freericks, B.K. Nikolic, and P. Miller. Superconductor-correlated metal-superconductor Josephson junctions: an optimized class for high speed digital electronics. *IEEE Trans. Appl. Supercond.*, 13(2):1089-1092, 2003.
- [40] V. Lacquaniti, S. Gonzini, S.Maggi, E. Monticone, R. Steni, and D. Andreone. Nb-based SNS junctions with Al and TaO barriers for a programmable Josephson voltage standard. *IEEE Trans. Appl. Supercond.*, 9(2):4245-4248, 1999.
- [41] A. B. Kaul, S.R. Whiteley, T. Van Duzer, Lei Yu, N. Newman, and J. Rowell. Internally shunted sputtered nbn Josephson junctions with a Tax barrier for nonlatching logic applications. *Appl. Phys. Lett.*, 78(1):99-101, 2001.
- [42] A. S. Barrera and M. R. Beasley. High-resistance SNS sandwich-type Josephson junctions. *IEEE Trans. Magn.*, MAG-23(2):866-868, 1987.
- [43] B. Baek, P. D. Dresselhaus, and S. P. Benz. Co-sputtered amorphous Nb-Si barriers for Josephson-junction circuits. *IEEE Trans. Appl. Supercond.*, 16(4):1966-1970, 2006.
- [44] Baek B, Dresselhaus PD, Benz SP. Thermal stability of Nb/a-NbxSi1-x/Nb Josephson junctions. *Phys. Rev. B* 2007, 75(5): 054514-9.
- [45] V. Lacquaniti, S. Gonzini, S.Maggi, E. Monticone, R. Steni, and D. Andreone, New Barriers for Fast-Switching SNS Josephson Junctions, *International Journal of Modern Physics B (IJMPB)*, 13(9-10), 1999.



- [46] Y. Chong, P. D. Dresselhaus, S. P. Benz, and J. E. Bonevich. Effects of Interlayer Electrode Thickness in Nb/(MoSi/Nb) Stacked Josephson Junctions. *Appl. Phys. Lett.*, 82(15):2467-2469, 2003.
- [47] A.S. Katkov, A.M. and Klushin, G.P. Telitchenko, R. Behr, J. Niemeyer, J., Challenges of Josephson junction arrays for ac voltage generation by microwave pulse power modulation, *IEEE Trans. Appl. Supercond.*, 15, 2, 352-355, 2005.
- [48] O. V. Karpov, V. M. Buchstaber, S. V. Sherstobitov, and S. I. Tertychniy, Josephson junction parameters for alternating current wave form synthesis, *J. Appl. Phys.* 100, 093907 (2006); DOI:10.1063/1.2358817, (2006).
- [49] Process RSFQ1D-1.2 Rapid Single Flux Quantum (RSFQ) - Design Rules for Nb/Al<sub>2</sub>O<sub>3</sub>-Al/Nb-Process, IPHT RSFQ.
- [50] M. Maezawa, V. Hirayama 10-bit rapid single flux quantum digital-to-analog converter for AC voltage standard *Journal of Physics: Conference Series*, 97(2008) 012161
- [51] Balashov, D. and Khabipov, M. and Buchholz, F.-I. and Niemeyer, J. SINIS process development for integrated circuits with characteristic voltages exceeding 250 uV, *IEEE Transactions on Applied Superc.*, 11, 1, 1070-1073, (2001).
- [52] Olaya, D. Baek, B. Dresselhaus, P.D. Benz, S.P. High-Speed Nb/Nb-Si/Nb Josephson Junctions for Superconductive Digital Electronics, *IEEE Transactions on Applied Superc.* 18, 4, 1797-1800, (2008).
- [53] V. Lacquaniti, C. Cagliero, S. Maggi, and R. Steni. Overdamped Nb/Al-AIO/Nb Josephson Junctions. *Appl. Phys. Lett.*, 86(4):042501, 2005.
- [54] V. Lacquaniti, N. De Leo, M. Fretto, A. Sosso, and M.Belogolovskii, Nb/Al-AIO<sub>x</sub>-Nb superconducting heterostructures: A promising class of self-shunted Josephson junctions, *J. Appl. Phys.*, 108(9): 108-115, 2010.
- [55] K. M. Schep and G. E. W. Bauer, *Phys. Rev. Lett.* 78, 3015, 1997; *Phys. Rev. B* 56, 15860, 1997.
- [56] D. Averin and A. Bardas, *Phys. Rev. Lett.* 75:1831, 1995.
- [57] J. Lee, J.Schurr, J. Nissilä, L.Palafox, and Ralf Behr, The Josephson two-terminal-pair impedance bridge, *Metrologia*, 47 45, 2010.
- [58] S. Benz et al An electronic measurement of the Boltzmann constant, *Metrologia*, 48 142, 2011.
- [59] S.A. Cybart, K. Chen, Y. Cui, Q. Li, XX Xi, and RC Dynes. Planar MgB<sub>2</sub> josephson junctions and series arrays via nanolithography and ion damage. *Appl. Phys. Lett.*, 88:012509, 2006.
- [60] A.M. Klushin, M. He, S.L. Yan, and N. Klein. Arrays of high-T Josephson junctions in open millimeter wave resonators. *Appl. Phys. Lett.*, 89:232505, 2006.
- [61] S.I. Borovitskii, A.M. Klushin, T.B Korotina, A.E. Pariiskii, S.K. Khorshev, Increasing the working voltage across chains of Josephson junctions, *Sov. Tech. Phys. Lett. (Engl. Transl)*, 11(6), 169, 1985
- [62] V. Lacquaniti, N. De Leo, M. Fretto, A. Sosso, F. Muller, J. Kohlmann, 1 V Programmable Voltage Standards based on SNIS Josephson Junctions Series Arrays, *Superc. Sci. Tech.*, 24: 045004-045007, 2011.
- [63] V. Lacquaniti, D. Andreone, N. De Leo, M. Fretto, S.Maggi, A. Sosso, and M.Belogolovskii. Analysis of the Temperature Stability of Overdamped Nb/Al-AIO/Nb Josephson Junctions, *IEEE Trans. Appl. Supercond.*, 17(2):609-612, June 2007.
- [64] V. Lacquaniti, D. Andreone, N. De Leo, M. Fretto, A. Sosso, and M.Belogolovskii, Engineering overdamped niobium-based Josephson junctions operating above 4.2 K, *IEEE Trans. Appl. Supercond.*, 19: 234-237, 2009.
- [65] V. Lacquaniti, N. De Leo, M. Fretto, S.Maggi, A. Sosso, Nb/Al/AIO<sub>x</sub>/Nb Overdamped Josephson Junctions above 4.2 K for Voltage metrology, *Appl. Phys. Lett.* , vol. 91; p. 252505 1-252505 3, 2007.

# Measurement of Harmonic Voltage, Current, and Power at Industrial Frequency

Lu Zuliang  
*National Institute of Metrology*  
*P. R. China*

## 1. Introduction

In recent years, power quality analyzers and revenue meters with harmonic analyzing function have increasingly been used to monitor the state of the power supply network. The precise capability of revenue meters to measure electrical energy under non-sinusoidal conditions has been questioned on the basis of some IEC [IEC, 2003] standards and international recommendations. For EMC, harmonic tests are prescribed in the IEC documents [IEC, 2001, 2002] to ensure that appliances including televisions, computers, lamps, and motors do not produce harmonic currents that exceed allowable limits when they are operated under sinusoidal voltage. The harmonic measurement is an important task too in the background of the smart grid.

Many harmonic analyzers with variable levels and harmonic sources with determined accuracies have been offered in the market even as new instruments are being developed. At the same time, these have created a need for national laboratories to provide calibration facilities and traceability for harmonic quantities, including harmonic voltage, harmonic current, and harmonic power.

To respond to the changing needs of the industry for non-traditional electrical measurements, many national laboratories (NMIs) have developed standards for power measurements under non-sinusoidal waveform conditions. A harmonic power standard equipment was also developed at the National Institute of Metrology of China (NIM) in 2006. This equipment can accurately calibrate harmonic power analyzers and harmonic sources, and calibrate the harmonic-analyzing-function of the revenue meters, including its measuring capability for electrical energy under non-sinusoidal conditions. The uncertainties of the standard equipment are less than  $30 \mu\text{V}/\text{V}$ ,  $36 \mu\text{A}/\text{A}$ , and  $42 \mu\text{W}/\text{VA}$  for harmonic voltage, harmonic current, and harmonic power measurements, respectively, at fundamental frequencies of 50 and 60 Hz and operating ranges of up to 50 A, 500 V, and 60<sup>th</sup> order harmonics.

The harmonic power standard equipment is introduced in this chapter as an example. Its measurement functions for harmonic voltage, harmonic current, and harmonic power at industrial frequency are described, with aspects such as its principle, setup, uncertainty evaluation, and experimental validation also covered.

The digital sampling technique is widely adopted for commercial harmonic analyzers. For standard equipment, this technique must be developed to a higher level to satisfy calibration requirements. The standard equipment of the NIM is also based on digital sampling techniques. Unlike that proposed in general sampling theory, however, the synchronization between sampling rate and signal frequency is not a stringent requirement for the NIM equipment. The leakage effect that results from such an asynchronous case is compensated for by the introduction of a novel algorithm. The engineering practicality of this algorithm is demonstrated and its calculation is limited to a couple of seconds, without the need for an especially large computing space.

The uncertainty evaluation for the standard equipment is based primarily on experiments. The standard equipment is tested using the national AC voltage and AC current standards at variable frequency points to determine frequency characteristics. Some special factors for harmonic conditions are then considered; these include the small harmonic components, leakage between harmonics, noise, and nonlinear effect of frequency. Finally, some experiments are designed and implemented to validate the uncertainty.

The principles and methods of the algorithm are designed to achieve higher accuracy. The uncertainty evaluation is carried out from the frequency to the harmonic feature. These attributes would be of significant reference for researchers, engineers, and students in developing higher quality commercial instruments or in general study. Based on these principles and methods the readers can improve system and reach a higher accuracy and a better function.

Further investigation to extend the capability of the algorithm for impedance measurement, higher frequency measurement, or other AC measurements would also be valuable.

The equipment system is introduced in Section 2. The algorithm is described in Section 3. The uncertainty of harmonic measurement is conceptually shown in Section 4. The harmonic voltage and harmonic current, their phase shifts, and harmonic power are discussed in Sections 5, 6, 7, and 8, respectively. The experimental validation is presented in Section 9.

The core materials on which this chapter is based are taken from [Lu et al., 2010], and some concepts in [Lu et al., 2008a, 2008b, 2008c] are extended. A primary form of the algorithm is discussed in [Lu, 1988], and its detailed analysis and application can be found in [Lu, 1991].

## 2. System

A block diagram of the NIM harmonic power standard equipment, including some general hardware, is shown in Fig. 2.1. Two commercial high-accuracy digital sampling voltmeters (DVMs) with type of HP3458A are employed as A/D converters (ADCs) to measure the instantaneous values of voltage and current signals. Non-sinusoidal signals are provided by two commercial programmable signal generators, one with the type of Fluke 6100A and another NST3500 from Chinese manufacturer. A set of resistive dividers have been developed to extend voltage ranges to 8, 15, 30, 60, 120, 240, and 500 V using selected resistors with a low time constant and low temperature coefficient. The resistance values are designed in such a way that the operating current does not introduce a significant heating effect. Another set of resistive dividers provide protection to reduce the effect of stray capacitance. A set of shunts (provided by SP of Sweden) are used to extend current ranges to 0.1, 0.2, 0.5, 1, 2, 5, 10, and 20 A, while a current transformer is adopted for 50 A.

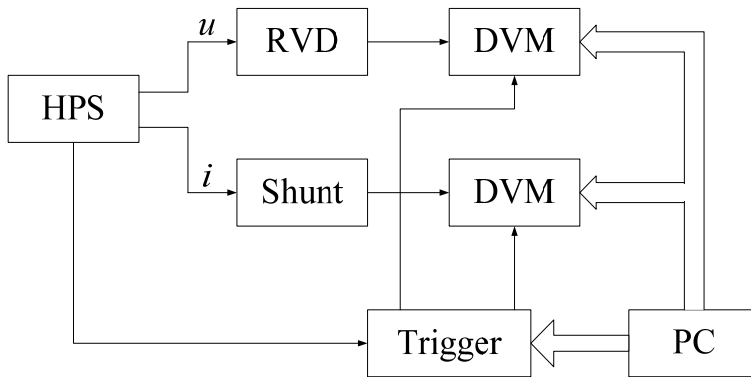


Fig. 2.1. Block Diagram of the harmonic power standard equipment; where HPS is harmonic power source, RVD is resistive voltage divider, and DVM is digital voltage meter.

The full-scale output signals of the voltage dividers and shunts are 0.8 V regardless of the voltage/current ranges, which are the input signals of the two corresponding DVMs.

A sampling trigger is developed to enable equal interval sampling between two contiguous sampling data, and synchronization sampling between voltage and current signals.

For the practical sampling strategy, 1680 samples over about four signal periods are computed on a PC, in accordance with the algorithm. The calculated results are voltage, current, and power at the DC, fundamental, and every harmonic order up to the 60<sup>th</sup>.

### 3. Algorithm

Generally, power sources and digital sampling meters have their own internal time clocks. Ensuring synchronous sampling capability in both instruments is difficult and often impossible. Such an asynchronous condition generates leakage errors in harmonic analysis because of the noncorrespondence between the accumulated values of the sampling data and the integral values in a precise period.

To resolve this issue, the authors in [Ihlenfeld et al., 2003] developed a method in which one time clock is adopted in the standard equipment, not only for the source (source of standard equipment, SSE), but also for the meter (meter of standard equipment, MSE). When a meter under test (meter under test, MUT) is calibrated, the SSE provides output signals for the MSE and MUT, resulting in synchronous sampling and appropriate calibration results. However, problems occur when calibrating another source (source under test, SUT) that has its own time clock. In such a case, due to two different time clocks are used, MSE calibration under synchronous sampling is difficult to accomplish, consequently resulting in poor calibration.

Non-integer-period sampling (NIPS) resolves this problem [Lu,1991]. It generates good results for both the MUT and SUT, without needing any special time clock and other equipment. The leakage effect persists, but can be overcome by the algorithm developed in this study for more general-purpose applications. The following section describes the algorithm.

### 3.1 Non-integer-period sampling

A period signal can be written as

$$y(x) = a_0 + \sum_{k=1}^{\infty} a_k \sin kx + b_k \cos kx, \quad (3.1.1)$$

where  $x = 2\pi t / \tau$ ,  $\tau$  is the signal period or fundamental period, and  $k$  denotes the harmonic order.

All cases can be expressed as

$$(n+\Delta)h = 2\pi m, \quad (3.1.2)$$

and the sampling data become

$$y_j = a_0 + \sum_{k=1}^{\infty} a_k \sin jkh + b_k \cos jkh$$

$$j = 0, 1, 2, \dots, n, \quad (3.1.3)$$

where  $h$  is the sampling interval,  $n$  denotes the number of samples in  $m$  periods, and  $\Delta$  represents a small quantity with the same means as  $n$ , noted as a deviation degree of the non-integer-period.

A sampling will be entitled as integer-period sampling (IPS) when  $\Delta = 0$ , and quasi-integer-period sampling (QIPS) when  $|\Delta| < 1$ . Generally, it will be expressed as  $|\Delta h| < \pi$ ; that is, NIPS.

Fig. 3.1 explains this concept, in which three periods of a signal is divided by 13 samples at the same interval. For case (a), an IPS sample is obtained when 13 samples are computed; for case (b), an NIPS sample is derived when 7 samples are computed; for case (c), a QIPS sample is obtained when 4 samples are calculated.

The key point lies in how many samples are obtained or treated. At any specified uncertainty, a certain IPS sample can always be found for any NIPS sample when the sample number is not limited, so that the NIPS sample is a part of the IPS sample [Lu, 1991].

Similar QIPS samples also exist, as shown in Fig. 3.1. These are the groups of five, eight, and nine samples.

### 3.2 Orthogonal check of the trigonometric function in NIPS

An accumulation operation is used in measuring harmonic quantities from samples of periodic signals. It is denoted as a linear operator/calculator  $A$ .

A few detailed forms of  $A$  were introduced in [Lu, 1988]. These are the trapezoid formula

$$Ay_j = \frac{1}{n} \left( 0.5y_0 + \sum_{j=1}^{n-1} y_j + 0.5y_n \right), \quad (3.2.1)$$

rectangular (Stenbakken's) compensation

$$Ay_j = \frac{1}{n + \Delta} \left( \sum_{j=0}^{n-1} y_j + \Delta y_n \right), \tag{3.2.2}$$

trapezoid compensation

$$Ay_j = \frac{1}{n + \Delta} \left[ \sum_{j=1}^{n-1} y_j + 0.5(1 + \Delta)(y_0 + y_n) \right]. \tag{3.2.3}$$

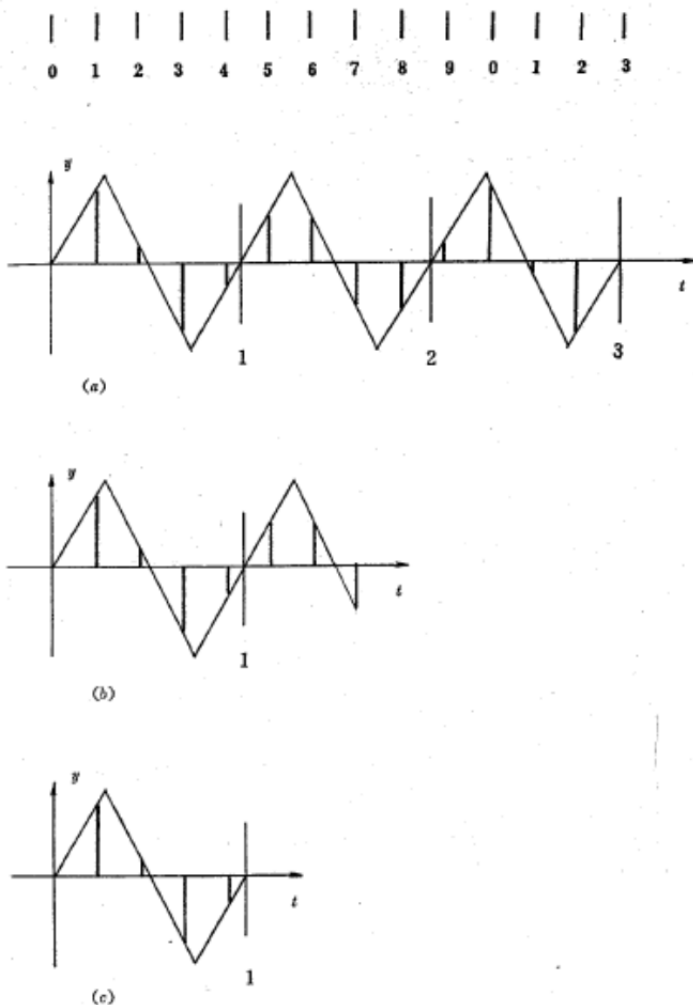


Fig. 3.1. (a) is an IPS sample, (b) is a NIPS sample, and (c) is a QIPS sample

In NIPS, the orthogonality of the trigonometric function appears as a deviation.

If

$$A \sin jkh = \alpha_k, \quad (3.2.4)$$

$$A \cos jkh = \beta_k. \quad (3.2.5)$$

Using the characteristic of the trigonometric function proves that

$$2A(\sin jkh \cos jlh) = \alpha_{k+l} + \alpha_{k-l}, \quad (3.2.6)$$

$$2A(\sin jkh \sin jlh) = -\beta_{k+l} + \beta_{k-l}, \quad (3.2.7)$$

$$2A(\cos jkh \cos jlh) = \beta_{k+l} + \beta_{k-l}. \quad (3.2.8)$$

In particular, when  $k = 0$ ,

$$\alpha_0 = A \sin j0h = A \cdot 0 = 0, \quad (3.2.9)$$

$$\beta_0 = A \cos j0h = A \cdot 1 = 1. \quad (3.2.10)$$

### 3.3 Fourier coefficient of the signal in NIPS

If the Fourier coefficients of the signal described in Eq. (3.1.1) are  $\hat{a}_0, \hat{a}_k, \hat{b}_k$ , the following hold:

$$\hat{a}_0 = a_0 + \sum_{k=1}^{\infty} \alpha_k a_k + \beta_k b_k, \quad (3.3.1)$$

$$\begin{aligned} \hat{a}_k &= 2\alpha_k a_0 + \sum_{l \neq k}^{1, \infty} (-\beta_{k+l} + \beta_{k-l}) a_l + (1 - \beta_{2k}) a_k, \\ &+ \sum_{l \neq k}^{1, \infty} (\alpha_{k+l} + \alpha_{k-l}) b_l + \alpha_{2k} b_k, \end{aligned} \quad (3.3.2)$$

$$\begin{aligned} \hat{b}_k &= 2\beta_k a_0 + \sum_{l \neq k}^{1, \infty} (\alpha_{k+l} - \alpha_{k-l}) a_l + \alpha_{2k} a_k, \\ &+ \sum_{l \neq k}^{1, \infty} (\beta_{k+l} + \beta_{k-l}) b_l + (1 + \beta_{2k}) b_k. \end{aligned} \quad (3.3.3)$$

If  $w$  is the maximum limit of the harmonic order of the signal shown in Eq. (3.1.1), the Fourier coefficients of the signal can be denoted by a vector  $\mathbf{a}$  with order  $2w+1$ :

$$\mathbf{a} = (a_0, a_1, a_2, \dots, a_w, b_1, b_2, \dots, b_w)^T. \quad (3.3.4)$$

Then, its DFT result can be expressed as another vector  $\hat{\mathbf{a}}$  :

$$\hat{\mathbf{a}} = (\hat{a}_0, \hat{a}_1, \hat{a}_2, \dots, \hat{a}_w, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_w)^T. \quad (3.3.5)$$

According to Eqs. (3.3.1), (3.3.2), and (3.3.3), we have

$$\hat{\mathbf{a}} = \mathbf{F}_R \mathbf{a}. \quad (3.3.6)$$

Thus, the results of  $\mathbf{a}$  can be calculated as follows:

$$\mathbf{a} = \mathbf{F}_R^{-1} \hat{\mathbf{a}}, \quad (3.3.7)$$

where  $\mathbf{F}_R$  is a matrix with order  $2w+1$ :

$$\mathbf{F}_R = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} & \mathbf{F}_{13} \\ \mathbf{F}_{21} & \mathbf{F}_{22} & \mathbf{F}_{23} \\ \mathbf{F}_{31} & \mathbf{F}_{32} & \mathbf{F}_{33} \end{bmatrix}. \quad (3.3.8)$$

Its detailed form will be dependent on the sampling rate and signal period, and will be independent of the amplitude and phase angle of the signal. That is, matrix  $\mathbf{F}_R$  is available for all signals when both the sampling rate and signal period are fixed.

Some interesting and useful characteristics of  $\mathbf{F}_R$  are as follows:

$$\mathbf{F}_{11} = 1$$

$$\mathbf{F}_{12} = \frac{1}{2}(\mathbf{F}_{21})^T = (a_1, a_2, \dots, a_w)$$

$$\mathbf{F}_{13} = \frac{1}{2}(\mathbf{F}_{31})^T = (\beta_1, \beta_2, \dots, \beta_w)$$

$$(\mathbf{F}_{22})_{kl} = (-\beta_{k+l} + \beta_{k-l})$$

$$(\mathbf{F}_{23})_{kl} = (a_{k+l} + a_{k-l})$$

$$(\mathbf{F}_{32})_{kl} = (a_{k+l} - a_{k-l})$$

$$(\mathbf{F}_{33})_{kl} = (\beta_{k+l} + \beta_{k-l})$$

where  $k, l = 1, 2, \dots, w$ .

When  $w = 3$ , the matrix of  $\mathbf{F}_R$  can be constructed as



1	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\beta_1$	$\beta_2$	$\beta_3$
$2\alpha_1$	$1-\beta_2$	$\beta_1-\beta_3$	$\beta_2-\beta_4$	$\alpha_2$	$\alpha_3-\alpha_1$	$\alpha_4-\alpha_2$
$2\alpha_2$	$\beta_1-\beta_3$	$1-\beta_4$	$\beta_1-\beta_5$	$\alpha_3+\alpha_1$	$\alpha_4$	$\alpha_5-\alpha_1$
$2\alpha_3$	$\beta_2-\beta_4$	$\beta_1-\beta_5$	$1-\beta_6$	$\alpha_4+\alpha_2$	$\alpha_5+\alpha_1$	$\alpha_6$
$2\beta_1$	$\alpha_2$	$\alpha_3+\alpha_1$	$\alpha_4+\alpha_2$	$1+\beta_2$	$\beta_1+\beta_3$	$\beta_2+\beta_4$
$2\beta_2$	$\alpha_3-\alpha_1$	$\alpha_4$	$\alpha_5+\alpha_1$	$\beta_1+\beta_3$	$1+\beta_4$	$\beta_1+\beta_5$
$2\beta_3$	$\alpha_4-\alpha_2$	$\alpha_5-\alpha_1$	$\alpha_6$	$\beta_2+\beta_4$	$\beta_1+\beta_5$	$1+\beta_6$

### 3.4 Determination of the period

In the algorithm, the period must first be determined. For the standard equipment, the signal period is cursorily known. A more accurate result can be calculated using  $\Delta$ .

The period end point can be assumed located between points  $y_n$  and  $y_{n+1}$ . Its value is exactly  $y_0$ . The period exists in the relationship of  $(y_n, y_0, y_{n+1})$ . The quantity is calculated [Lu, 1986] as

$$\Delta = (y_0 - y_n) / (-y_n + y_{n+1}) . \quad (3.4.1)$$

The other similar relationships of  $(y_0, y_{n+1}, y_1)$  can also be used. A more precise formula is [Zhang J.Q., 1996]

$$\Delta = (y_0 + y_1 - y_n - y_{n+1}) / (-y_0 + y_1 - y_n + y_{n+1}) . \quad (3.4.2)$$

The selection of the starting point of  $y_0$  is important; it must be located in a region with a large slope. Under sinusoidal conditions, for example, it should be a zero-crossing point, or a point with a minimum value among sampling data.

Other approaches to determining the period are also available. For example, the relationships of the calculated values of phase angles as (5.1.5) about the first period and the second period [Dai, 1989] can be used.

### 3.5 Practical implementation of DFT and compensation

For input sampling data, the practical implementing process of DFT is as follows:

1. to determine  $\Delta$  according to Eq. (3.4.1) or (3.4.2);
2. to calculate the sampling/discrete interval of  $h$  according to Eq. (3.1.2)
3. to calculate the base functions of  $\sin jkh$  and  $\cos jkh$  which are needed by DFT;
4. to implement DFT with the trapezoid compensation calculator  $A$  in Eq.(3.2.3).

Where an important point is that the base functions of  $\sin jkh$  and  $\cos jkh$  needed by DFT come from the practical case including the signal period and the sampling rate, but not come only from the sampling rate of ADC.

When application of Eq. (3.2.3) in DFT, the term of  $y_j$  shall be replaced by  $y_j \sin jkh$  or  $y_j \cos jkh, j = 0, 1, 2, \dots, n$ .

After this DFT a further compensation will be carried out for higher accuracy. For this aim, in the fact, the eqn. (3.3.7) is recognized by many researchers, but its calculation is very difficult. A large-capacity RAM and long computing time are required to calculate  $\mathbf{F}_R^{-1}$  because of the huge matrix. Its  $2w+1$  order, for example, is 121 when  $w = 60$  in the NIM standard equipment.

For practical implementation,  $\mathbf{F}_R$  must be simplified.

As mentioned in Section 3.3, only two types of quantities,  $a_k$  and  $\beta_k$ , are necessary for all the elements of the matrix. They should be expressed in analyzable forms. According to the expression of  $a_k$  and  $\beta_k$  in Eqs. (3.2.4) and (3.2.5), as well as the expression of trapezoid compensation operator  $A$  in Eq. (3.2.3), the quantities can be analyzed and described as

$$\alpha_k = \frac{1}{n + \Delta} \left( \sin^2 \frac{\Delta kh}{2} \operatorname{ctg} \frac{kh}{2} - \frac{\Delta}{2} \sin \Delta kh \right), \quad (3.5.1)$$

$$\beta_k = \frac{1}{n + \Delta} \left( -\frac{1}{2} \sin \Delta kh \operatorname{ctg} \frac{kh}{2} + \Delta \cos^2 \frac{\Delta kh}{2} \right). \quad (3.5.2)$$

However,  $\beta_k = 1$ ,  $a_k = 0$  when  $k = 0$ . Alternatively,  $a_{-k} = -a_k$ ,  $\beta_{-k} = \beta_k$ .

Some special controls are adopted in the NIM standard equipment to let  $\|\mathbf{X}\| \ll 1$ , where  $\mathbf{F}_R = \mathbf{I} + \mathbf{X}$  and  $\mathbf{I}$  is a unit matrix. Thus,  $\mathbf{F}_R^{-1} \approx \mathbf{I} - \mathbf{X}$ , where  $\mathbf{X}$  can be directly provided according to Eqs. (3.3.6), (3.5.1), and (3.5.2). The calculation of  $\mathbf{F}_R^{-1}$  becomes a very easy process. The author notes here that it will be not necessary for the hardware control in principle.

The precision result,  $\mathbf{a}$ , is directly calculated without any intermediate process. A complete calculation that includes DFT and compensation can be completed within a couple of seconds, without the necessity for an especially large computing space.

The compensation results are dependent on the ratio of the computational error/uncertainty of  $\Delta$  to  $n$  [Lu, 2008d].

### 3.6 Simulation and the effect of noise

A simulation test is designed to verify the effect of the algorithm and results, as evaluated against those derived from the general DFT method with no compensation. In the general method the sampling interval is regarded as  $2\pi m/n$ , but in the compensation method used in this study, it is  $2\pi m/(n+\Delta)$ , as shown in Eq. (3.1.2). However in this DFT this  $\Delta$  value is asked to calculated from the sampling data.

The sampling data from a sinusoidal signal with an amplitude of 0.8 V (to check for potential computational errors) are simulated on a computer with same interval of  $2\pi m/(n+\Delta)$  (it becomes a NIPS or QIPS precisely). The sampling rate is set at 60 sampling data in one period with different  $\Delta$  values. The general DFT and proposed algorithm are then used to handle the same sampling data. Their computing values for signal amplitude are provided and then compared with the set value (0.8 V) to reveal the relative errors (Table 3.6.1).

$\Delta$	Rel. error/( $\mu\text{V}/\text{V}$ )	
	General DFT	Proposed algorithm
0.5	3905	10.0
0.2	1626	1.9
0.1	822	0.8
0.05	413	0.4
0.02	166	0.1
0.01	83	0.1

Table 3.6.1. Simulation results for amplitude,  $y = 0.8\sin \omega t$  (V),  $n=60$

The results show that when  $n = 60$  and  $\Delta$  is controlled within 0.05, the magnitude of the compensation effect can be as much as 1000 times. The relative error reaches the  $10^{-7}$  level.

A similar simulation is conducted for the phase difference between two signals with the same sampling rate, but the phase difference is set at  $60^\circ$ . The results are shown in Table 3.6.2.

$\Delta$	Error/ $\mu\text{rad}$	
	General DFT	Proposed algorithm
0.5	-3714	-13
0.2	-1615	-2.7
0.1	-828	-1.1
0.05	-420	-0.5
0.02	-169	-0.2
0.01	-85	-0.1

Table 3.6.2. Simulation results for phase difference,  $n=60$

On the basis of this analysis, we design a practical sampling method in which 1680 samples over about 4 periods and a  $\Delta$  value controlled within 0.04 is applied in the equipment.

For a fixed  $\Delta$ , the phase-angle values in units of rad are different in view of varied harmonic components. At the fundamental, if the value is  $2\pi m\Delta$ , but at the  $k$ -th harmonics, it is  $2\pi km\Delta$ . Therefore, the same  $\Delta$  for higher harmonics produces different errors.

In simulating this case, the same sampling point construction is used for all the harmonics of the non-sinusoidal signal. This approach, which we call "harmonic discrete division," is the foundation of the investigation on harmonic sampling measurement.

This method is adopted to investigate the performance of the algorithm for harmonics. A practical condition is considered; i.e., 1680 samplers (with  $\Delta$ ) over 4 fundamental periods. When the 10<sup>th</sup> order harmonics is simulated, the 1680 points are constructed over 40 sinusoidal waveforms. The results are shown in Table 3.6.3.

Order	1	10	20	30	40	50	60
Error/(nV/V)	0.2	21	68	102	74	55	298

Table 3.6.3. Amplitude error after compensation for harmonic components,  $y = 0.8\sin \omega t$  (V) for every harmonic,  $\Delta = 0.04$ ,  $n=1680$

The errors that occur after compensation reaches the  $10^{-10}$  level for the fundamental, and are less than  $3 \times 10^{-7}$  up to the 60<sup>th</sup> order harmonics. This result indicates that the leakage caused by the non-integer-period has been resolved.

Many other tests, including some complex waveforms, are conducted. A characterizing waveform signal discussed below is tested, in which all harmonic components from DC to 60<sup>th</sup> order harmonic exist with a relative high percent of amplitude, the largest error is  $0.2 \times 10^{-6}$ . This result can be disregarded for the final estimated uncertainty.

The practical specifications are determined in experiments, such as the calibration traceable to the primary standard of the AC voltage and AC current of the NIM. In such cases, the noise becomes the primary determining factor for uncertainties.

The experimental results show that the noise effect for sinusoidal signal is about  $(0.5 \text{ to } 2.0) \times 10^{-6}$ , as validated by Monte Carlo test results [Xue, 2011]. This means that the error from the algorithm (less than  $1.0 \times 10^{-6}$ ) is smaller than the effect of noise, and additional efforts to derive higher level algorithm results are not necessary.

## 4. Concept for the uncertainty of harmonic measurement

### 4.1 Fiducial error and fiducial uncertainty

Relative error and relative uncertainty are two concepts often used in calibrations for general quantities. These concepts have also been used in many expressions of harmonic measurement. However, the example given in this study illustrates the need to revisit these concepts.

A harmonic analyzer is used to measure a non-sinusoidal voltage signal with a fundamental 100 V and a 5<sup>th</sup> order harmonic 10 V. It measures the 5<sup>th</sup> order harmonics as 10.0101 V, of which 10 mV is a leakage error from the fundamental (100 V), whereas 0.1 mV is an error only from the 5<sup>th</sup> order harmonic (10 V). In this case, the relative measurement error of the 5<sup>th</sup> order harmonic of the analyzer is  $10.1 \text{ mV}/10 \text{ V} = 0.1\%$ .

In another measurement, the analyzer measures a signal with 100 V of the fundamental plus 0.1 V of the 5<sup>th</sup> order harmonic. In this case, the leakage error from the fundamental (10 mV) persists because of the fact that the 100 V fundamental is unchanged, but the error that comes only from the 5<sup>th</sup> order harmonic may be 1  $\mu\text{V}$ . Hence, the relative measurement error is  $10 \text{ mV}/0.1\text{V} = 10\%$ , and not 0.1% as in the first case.

The example illustrates that the specifications of an analyzer are dependent on the magnitude of the signal to be measured for the relative error concept. Some error components may come from the fundamental. In such a case, the harmonic relative error of the analyzer is variable when the harmonic amplitude varies. This attribute does not correspond with that indicated in the typical error concept.

The fundamental is the main component in a general non-sinusoidal signal, whereas the harmonics are secondary components. Therefore, the selection of the measurement range in characterizing harmonic analyzers is dependent/determined primarily by the value of the fundamental component.

However, the concept of total harmonic distortion (THD) is related only to the fundamental component or the total RMS value of the harmonics including the fundamental, and not to the individual harmonic component itself.

If the fundamental component is used as reference, then the error is the same in both cases, as shown in the earlier example. That is,  $10 \text{ mV}/100\text{V} = 0.01\%$ . However, we introduce a new error concept called the *fiducial harmonic error*, developed by the International Vocabulary of Metrology (VIM, Second edition) [BIPM et al., 1993a] (regrettably, the term is not embodied in the Third edition, JCGM 200:2008). This concept is defined as follows:

$$\delta V_k = \Delta V_k / V_1, \quad (4.1.1)$$

where  $\Delta V_k$  is the absolute error of the  $k$ -th harmonic, and  $V_1$  is value of the fundamental.

The harmonic uncertainty concept is expanded with a relative uncertainty concept,  $u(\Delta V_k/V_k)$ , when the quantity itself,  $V_k$ , is taken as the reference and the absolute uncertainty is defined as  $u(\Delta V_k)$ :

$$u(\Delta V_k/V_k) = [u(\Delta V_k)]/V_k. \quad (4.1.2)$$

Changing the reference from  $V_k$  into  $V_1$ , we introduce a new harmonic uncertainty concept and denote it as  $u(\Delta V_k/V_1)$ :

$$u(\Delta V_k/V_1) = [u(\Delta V_k)]/V_1. \quad (4.1.3)$$

This harmonic uncertainty is defined as the *fiducial harmonic uncertainty*.

The relationship between the fiducial and relative harmonic uncertainty is

$$u(\Delta V_k/V_1) = u(\Delta V_k/V_k) \times (V_k/V_1). \quad (4.1.4)$$

The fiducial harmonic uncertainty is a reasonable choice in expressing harmonic measurement uncertainties, as explained previously.

## 4.2 Characterizing waveform signal

In Eq. (4.1.4), the ratio of  $V_k/V_1$  is an indeterminate variable. To estimate the uncertainty of the equipment, this ratio should be a fixed value to cover all the possible calibration cases. The IEC documents, international recommendations for electricity meters, and EMC provide some typical waveforms [IEC, 2001, 2002, 2003] that can be used to estimate these fixed ratio values. The National Research Council of Canada (NRC) proposed eight different reference waveforms based on actual field-recorded distorted waveforms that can be used for same aim [Arseneau et al., 1995a].

In conclusion, after considering all the different potential test waveforms that can be used, some rules can be assumed as follows:

$$V_0/V_1 \leq 0.5,$$

$$V_k/V_1 \leq 1, \text{ when } 2 \leq k \leq 10,$$

$$V_k/V_1 \leq \gamma/k, \text{ when } 10 < k \leq 60,$$

where  $V_k$  is the amplitude of the  $k$ -th harmonic component,  $V_0$  denotes the DC component,  $V_1$  represents the fundamental, and  $\gamma$  is a ratio factor.

A characterizing waveform signal, determined according to the aforementioned limit values, is proposed (Table 4.2.1). Its corresponding waveform is shown in Fig 4.2.1. This is only a theoretical signal ( $THD = 280\%$ ), conjured to estimate the uncertainty of the harmonic power standard equipment. The signal covers all the possible cases under calibration, so that the harmonic uncertainty claim has universality. It is frequently used for the general analysis and digital evaluation of uncertainty for the harmonic power standard of the NIM [Lu, 2010].

Order	$\gamma$	Ratio	Order	$\gamma$	Ratio	Order	$\gamma$	Ratio
0		0.500	21	3	0.143	41	1	0.024
1		1.000	22	3	0.136	42	1	0.024
2		0.900	23	3	0.130	43	1	0.023
3		0.900	24	3	0.125	44	1	0.023
4		0.900	25	3	0.120	45	1	0.022
5		0.900	26	3	0.115	46	1	0.022
6		0.800	27	3	0.111	47	1	0.021
7		0.800	28	3	0.107	48	1	0.021
8		0.800	29	3	0.103	49	1	0.020
9		0.800	30	3	0.100	50	1	0.020
10		0.750	31	2	0.065	51	1	0.020
11	8	0.727	32	2	0.063	52	1	0.019
12	7	0.583	33	2	0.061	53	1	0.019
13	6	0.462	34	2	0.059	54	1	0.019
14	5	0.357	35	2	0.057	55	1	0.018
15	4	0.267	36	2	0.056	56	1	0.018
16	3	0.188	37	2	0.054	57	1	0.018
17	3	0.176	38	2	0.053	58	1	0.017
18	3	0.167	39	2	0.051	59	1	0.017
19	3	0.158	40	2	0.050	60	1	0.017
20	3	0.150						

Table 4.2.1. Amplitude ratio of the characterizing signal,  $V_k/V_1$

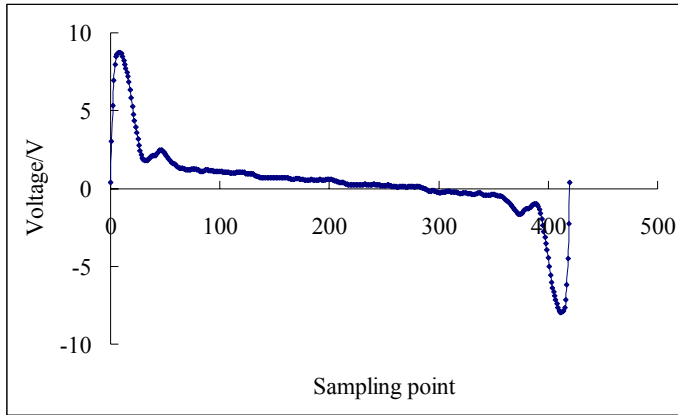


Fig. 4.2.1. Waveform of the characterizing signal.

## 5. Traceability of harmonic voltage and its uncertainty

### 5.1 Determination of harmonic voltage

When Eq. (3.1.1) expresses a non-sinusoidal voltage signal and Eq. (3.1.3) denotes the sampling data under the sampling model of Eq. (3.1.2), the DC voltage, fundamental voltage, and harmonic voltage components can be obtained after calculation according to Eq. (3.3.7). These are expressed as follows:

$$V_0 = a_0, \quad (5.1.1)$$

$$V_1 = \sqrt{a_1^2 + b_1^2}, \quad (5.1.2)$$

$$\varphi_{V1} = \arctan(b_1 / a_1), \quad (5.1.3)$$

$$V_k = \sqrt{a_k^2 + b_k^2}, \quad (5.1.4)$$

$$\varphi_{Vk} = \arctan(b_k / a_k), \quad (5.1.5)$$

where  $\varphi_{V1}$ ,  $\varphi_{Vk}$  are the phase angles (against a reference time point) of the fundamental and harmonics, respectively, and  $k = 2, 3, \dots, w$ .

The experimental determination of the technical specifications of the NIM harmonic power standard equipment is discussed next, beginning with harmonic voltage.

### 5.2 Frequency characteristic

The frequency characteristic of the standard equipment can be initially determined by adopting the existing National AC Voltage Primary Standard for thermo-converters. In this process, the stable sinusoidal signals are exported at the selected frequency points that correspond to the fundamental, 10<sup>th</sup>, 20<sup>th</sup>, 30<sup>th</sup>, 40<sup>th</sup>, 50<sup>th</sup>, and 60<sup>th</sup> orders. Their amplitudes

are set to correspond to the full range. Such signals are connected in parallel to the AC Voltage Standard and the harmonic power standard equipment (DVM or/and RVD). The experimental results show the differences than the AC Voltage Standard and its uncertainties.

This frequency characteristic is a special frequency attribute. The principal difference is that in the sampling measurement for high frequencies, harmonic discrete division is adopted (see Section 3.6). A similar case exists for the current discussed in Section 5.1 and the phase shift mentioned in Section 7.4, but this is not discussed for simplification.

### 5.2.1 Digital Voltage Meter

Two DVMs (one for voltage, and another for current) are tested. The input range of 0.8 V is set as the test voltage. The DVMs are self-calibrated in advance through short input operation, and are connected to the reference voltage.

The results indicate that both DVMs have similar characteristics. The relative differences/errors ( $\Delta V_k/V_k$ ) are

less than 1  $\mu\text{V}/\text{V}$  at the fundamental ,  
 less than 10  $\mu\text{V}/\text{V}$  within the 50<sup>th</sup> order,  
 less than 20  $\mu\text{V}/\text{V}$  at the 60<sup>th</sup> order.

Their standard deviations are all less than 3  $\mu\text{V}/\text{V}$ .

### 5.2.2 Resistive Voltage Divider

In this experiment, test voltages of 50, 100, 200, and 500 V are set. The resistive voltage dividers and buffers are tested together with the DVM. The relative differences/errors ( $\Delta V_k/V_k$ ) in the results are not considered for correction, but are taken as a part of uncertainty according to GUM [BIPM et al., 1995b], so that relevant correction will not be necessary in future simplified practical operations. After evaluating these results and their standard deviations, the uncertainties of the resistive voltage divider (RVD) are obtained (Table 5.2.1).

Order	1	10	20	30	40	50	60
50 V	5.22	8.05	7.36	5.25	12.6	14.4	36.3
100 V	8.45	9.95	10.6	10.0	8.83	6.58	35.6
200 V	7.68	8.53	7.86	10.8	15.1	19.2	47.6
500 V	9.50	8.81	12.6	19.6	29.2	26.7	64.2

Table 5.2.1. Relative uncertainties of the voltage without correction,  $u(\Delta V_k/V_k)/(\mu\text{V}/\text{V})$  ( $k = 1$ )

## 5.3 Harmonic characteristic

The frequency characteristic of the standard equipment obtained in the section above is not the harmonic characteristic required by harmonic measurement. Some other factors that focus on the fiducial uncertainty in Eq. (4.1.3) are considered.



### 5.3.1 Effect of the small component

The harmonic components are smaller in practical non-sinusoidal signals. The uncertainty estimation at full range in Table 5.2.1 cannot be directly applied to smaller amplitudes. A general method for non-linear amplitudes can be employed, where the ratio of  $(V_k/V_1)$  is replaced by  $[0.8(V_k/V_1)+0.2]$ , in which a minimum value of 0.2 (or 20%) is maintained even if  $V_k = 0$ . The ratio 0.8:0.2 (or 80%:20%) is taken from the experimental results. Eq. (4.1.4) then becomes

$$u(\Delta V_k/V_1) = [0.8(V_k/V_1) + 0.2] \times u(\Delta V_k/V_k). \quad (5.3.1)$$

### 5.3.2 Leakage and noise

When a frequency signal (for example, a sinusoidal signal corresponding to the 10<sup>th</sup> order harmonics) is tested in the single tone experiment, other components (from the DC to the 60<sup>th</sup> order) aside from the input signal itself appear in the DFT computational results. These components should be zero but are non-zero in practice possibly because the output of the source is not a pure single tone but with side bands; this result may also be attributed to both leakage due to faulty compensation and noise.

To overcome the side bands of the source, 3 sets of data (7<sup>th</sup>, 8<sup>th</sup>, 9<sup>th</sup>) ahead of the input signal (10<sup>th</sup>) and 3 sets of data behind (11<sup>th</sup>, 12<sup>th</sup>, 13<sup>th</sup>) that same signal in the DFT computational results are excluded in the analysis below.

The rest of the data reflect mainly leakage and noise. An “enlargement test” is designed and implemented to distinguish between both parts. The effect of the non-integer-period is enlarged 10 times ( $\Delta$  is enlarged from 0.04 to 0.4). The difference between the two groups of data is regarded as the influence from leakage only, and the remainder is regarded as noise.

The error of the  $k$ -th harmonic includes all leakages from the DC to the 60<sup>th</sup> order aside from itself. Therefore, all the contributions from the other harmonics are accumulated into the uncertainty of the  $k$ -th harmonic.

The contribution of the leakage from the experimental results is  $2.5 \mu\text{V}/V$ , and the noise dependent on frequency is  $6.0 \mu\text{V}/(k = 1)$ .

### 5.3.3 Non-linearity effect

When two sinusoidal signals are simultaneously input, such as the 1<sup>st</sup> and 5<sup>th</sup> orders, and when the input circuit of the meter (or the output circuit of the source) exhibits non-linearity of frequencies, two new sinusoidal signals, the 4<sup>th</sup> and 6<sup>th</sup> orders, appear in the results. This effect occurs too in the single tone experiment.

We design and implement an experiment, in which a signal with the fundamental and 5<sup>th</sup> order is measured. Another signal with only the 5<sup>th</sup> order is measured. The difference between the two above-mentioned DFT results at the 4<sup>th</sup> and 6<sup>th</sup> orders reflects this non-linearity. According to the experimental results, the contribution of this effect is  $2.2 \mu\text{V}/V$  ( $k = 1$ ), including other possible influences from the signal source.

## 5.4 Conclusion

The small component effect of Eq. (5.3.1) is a part of the uncertainty. This component is called the *traceability component* because of its dependence on the voltage unit.

The leakage, noise, and non-linearity are other parts, but they are independent of the voltage unit. They can be regarded as the *resolution component*, expressed as SQR  $(2.5^2 + 6.0^2 + 2.2^2) = 6.85 \mu\text{V}/\text{V}$  ( $k = 1$ ) on the basis of the results in Section 5.3.

Considering the two components (traceability and resolution), the fiducial uncertainty of Eq. (5.3.1) becomes

$$u_c^2(\Delta V_k/V_1) = [0.8(V_k/V_1) + 0.2]^2 u^2(\Delta V_k/V_k) + 6.85^2, \quad (5.4.1)$$

where the parameters of  $u(\Delta V_k/V_k)$  can be introduced from Table 5.2.1, and the ratio of  $(V_k/V_1)$  can be adopted from the characterizing signal in Table 4.2.1. The final uncertainties are calculated and shown in Table 5.4.1.

Order	1	10	20	30	40	50	60
50 V	8.6	9.4	7.2	7.0	7.5	7.5	10.3
100 V	10.9	10.5	7.6	7.4	7.2	7.0	10.2
200 V	10.3	9.7	7.3	7.5	7.7	8.0	12.3
500 V	11.7	9.8	7.9	8.8	9.8	8.9	15.3

Table 5.4.1. Fiducial uncertainty under the characterizing signal condition,  $u_c(\Delta V_k/V_1)/(\mu\text{V}/\text{V})$  ( $k = 1$ )

An estimation of  $30 \mu\text{V}/\text{V}$  ( $k = 2$ ) for all the voltage ranges can be obtained. At 100 V, however, this estimation amounts to  $20 \mu\text{V}/\text{V}$ .

## 6. Traceability of harmonic current and its uncertainty

### 6.1 Determination of harmonic current

When Eq. (3.1.1) expresses a non-sinusoidal current signal, and Eq. (3.1.3) denotes the sampling data under the sampling model of Eq. (3.1.2), after computation according to Eq. (3.3.7), the DC current, fundamental current, and harmonic current components can be obtained thus:

$$I_0 = a_0, \quad (6.1.1)$$

$$I_1 = \sqrt{a_1^2 + b_1^2}, \quad (6.1.2)$$

$$\varphi_{11} = \arctan(b_1 / a_1), \quad (6.1.3)$$

$$I_k = \sqrt{a_k^2 + b_k^2}, \quad (6.1.4)$$

$$\varphi_{1k} = \arctan(b_k / a_k), \quad (6.1.5)$$

where  $\varphi_{1l}$ ,  $\varphi_{lk}$  are the phase angles (against a reference time point) of the fundamental and harmonics, respectively, and  $k = 2, 3, \dots, w$ .

Similar experiments and computations are carried out for the harmonic current specifications of the NIM harmonic power standard equipment.

### 6.2 Frequency characteristic

The shunts in the equipment are tested. The relative differences/errors are not considered as corrections. The current traceability uncertainty in the full range of every shunt is estimated and shown in Table 6.2.1, including the contribution from the Primary Standards (DC resistance, AC/DC transfer, and DC voltage).

Order	1	10	20	30	40	50	60
$\leq 5A$	10.6	11.8	11.8	11.7	10.3	11.3	17.0
10 A, 20 A	14.6	15.5	19.1	19.0	22.5	27.4	34.5

Table 6.2.1. Traceability uncertainty in the full range of shunts without correction,  $u(\Delta I_k/I_k)/(\mu A/A)$  ( $k = 1$ )

An experiment is designed and performed to check the aforementioned estimation (Fig. 6.2.1). An inductive shunt is developed [Zhang J.T. et al., 2007] to produce two equal currents with very high accuracy. In a general series connection of two resistors, a potential at the middle point causes error. The defect is addressed in this study. The results obtained satisfy the estimation in Table 6.2.1.

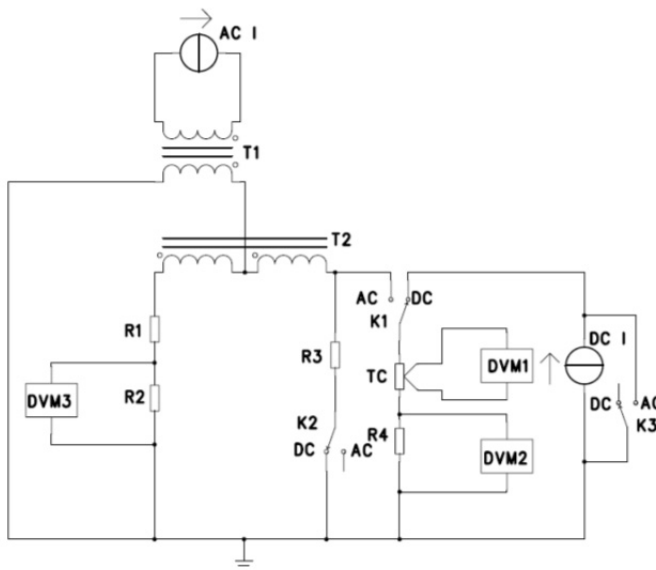


Fig. 6.2.1. Circuit of current traceability.

### 6.3 Harmonic characteristic

Similar to the voltage case, the *fiducial uncertainty of harmonic current*  $u(\Delta I_k/I_1)$  is obtained as

$$u_c^2(\Delta I_k/I_1) = u^2(\Delta I_k/I_1) + 7.28^2, \quad (6.3.1)$$

where

$$u(\Delta I_k/I_1) = [0.7(I_k/I_1) + 0.3] u(\Delta I_k/I_k) \quad (6.3.2)$$

is the traceability component in which non-linearity of amplitude is considered. For the resolution component, 2.5, 6.0, and 3.3  $\mu\text{A}/\text{A}$  are experimentally determined for leakage, noise, and non-linearity of frequencies, respectively.

The digital computational results for the characteristic signal condition are expressed in Table 6.3.1.

Order	1	10	20	30	40	50	60
$\leq 5$ A	14	14	11	11	10	10	11
10 A, 20 A	18	16	12	12	12	13	15
50 A	18	16	12	12	12	13	15

Table 6.3.1. Combined fiducial uncertainty in the characteristic signal,  $u_c(\Delta I_k/I_1)/(\mu\text{A}/\text{A})$  ( $k = 1$ )

The table shows that the experimental results for 50 A are derived from the current transformer. It also shows that the maximum result is 36  $\mu\text{A}/\text{A}$  ( $k = 2$ ), which can be used as an estimation for the harmonic measurement of the standard equipment.

## 7. Determination of phase shift and its uncertainty

The phase difference between voltage and current is an important quantity in power measurement. In harmonic power measurement, every phase difference between the harmonic voltage and harmonic current at the same order from the 2<sup>nd</sup> to the 60<sup>th</sup> must be determined. Two problems are discussed: how the phase difference is measured and how its uncertainty is evaluated.

### 7.1 Measurement of the phase difference of two signals

#### 7.1.1 Phase difference of two voltage signals

Phase difference can be measured using the sampling approach. The phase difference between two voltages can be calculated using the DFT sampling data results [Lu et al., 2006]:

$$\varphi_k = \arctan(b_{2k}/a_{2k}) - \arctan(b_{1k}/a_{1k}), \quad (7.1.1)$$

where  $\varphi_k$  is the phase difference of the  $k$ -th harmonics between voltages  $y_2$  and  $y_1$ :

$$y_1 = a_{10} + \sum_{k=1}^w (a_{1k} \sin k\omega t + b_{1k} \cos k\omega t)$$

$$y_2 = a_{20} + \sum_{k=1}^w (a_{2k} \sin k\omega t + b_{2k} \cos k\omega t)$$

$a_{1k}, b_{1k} (a_{2k}, b_{2k})$  are the Fourier coefficients of  $y_1 (y_2)$ .

The method involves sampling the two voltage signals simultaneously using two ADCs (i.e., DVMs in the standard equipment), and applying DFT to the sampling data and then computing phase difference according to Eq. (7.1.1).

The algorithm introduced in Section 3 can be applied here for precise Fourier coefficients to result in precise phase difference.

Because two different ADCs are used, an intrinsic phase difference occurs, thereby affecting the results. A voltage can be connected parallel to the two channels and the above-mentioned method can be implemented to determine this intrinsic error, which can then be corrected in succeeding measurements [Svensson,1998]. This process can be called **producing a 0° standard**.

### 7.1.2 Uncertainty of phase difference

The phase angle definition is included in Eq. (7.1.1); that is,

$$\varphi = \arctan(b/a). \quad (7.1.2)$$

Therefore, its uncertainty can be expressed by the amplitude

$$u^2(\varphi) = u^2(\arctan(b/a)) = [b u(a)/(a^2 + b^2)]^2 + [a u(b)/(a^2 + b^2)]^2.$$

If  $u(a) \geq u(b)$ ,

$$u(\varphi) \leq u(a)/\sqrt{a^2 + b^2} = u(a)/c, \quad (7.1.3)$$

where  $u(a)$  is an absolute uncertainty, whose concept is similar to absolute uncertainty  $u(\Delta V_k)$  in Section 4.

In general,  $a = (1 \sim 0.71) c \leq c$ ; thus,

$$u(\varphi_{V_k}) = (1 \sim 0.71) u(\Delta V_k)/V_k \leq u(\Delta V_k)/V_k = u(\Delta V_k/V_k).$$

The uncertainty of phase difference  $u(\Delta\varphi_{V_k})$  between two voltages can be expressed as

$$u^2(\Delta\varphi_{V_k}) = u^2(\Delta V_{1k}/V_{1k}) + u^2(\Delta V_{2k}/V_{2k}). \quad (7.1.4)$$

The relative uncertainty of  $u(\Delta V_k/V_k)$  is discussed in Section 5 for voltage (and Section 6 for current), and determined in experiments. However, some uncertainty factors can be discussed further:

1. Given that the phase angle is dependent on the ratio of Fourier coefficients, the uncertainty of the voltage standard from traceability has no function here.
2. The standard deviation of measurement is an important factor.
3. Leakage and noise are also factors.

4. The starting point of sampling brings about variation in Fourier coefficients  $a$  and  $b$ . However, the experiment shows that their influence can be disregarded.
5. The compensation of the algorithms is also a primary factor. However, the experiment proves that the residual effect occurring after the compensation using Eq. (3.3.7) is less than the noise. This effect can also be disregarded.

**7.1.3 Phase difference between a voltage signal and a current signal**

Only the shunt in the equipment is discussed in this section. The input signal is a current and the output signal is a voltage, shown in its equivalent circuit in Fig. 7.1.1, wherein points A and B denote the current input terminal, and C and D represent the voltage output terminal.  $R$  denotes resistance,  $L$  represents inductance, and  $C$  denotes capacitance. When the input current is taken as reference, the output voltage has a phase shift  $\varphi$ :

$$\varphi \approx \omega L / R - \omega RC . \tag{7.1.5}$$

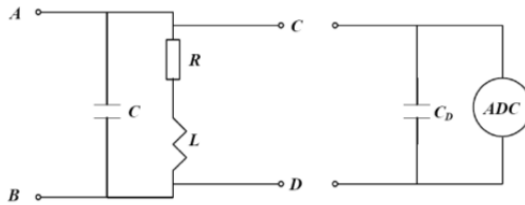


Fig. 7.1.1. Equivalent circuit of shunts.

When an ADC is used to measure the voltage, its input capacitor  $C_D$  may generate a new measurement result:

$$\varphi \approx \omega L / R - \omega RC - \omega RC_D , \tag{7.1.6}$$

where the function of  $C_D$  cannot be disregarded. To overcome such influencing factors, a substitution method-based sampling measurement is developed [Wang et al., 2008].

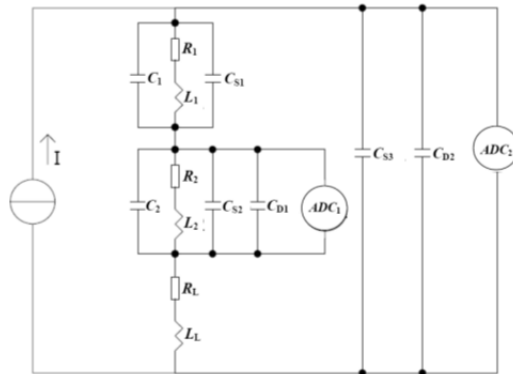


Fig. 7.1.2A. Phase shift measurement step 1.

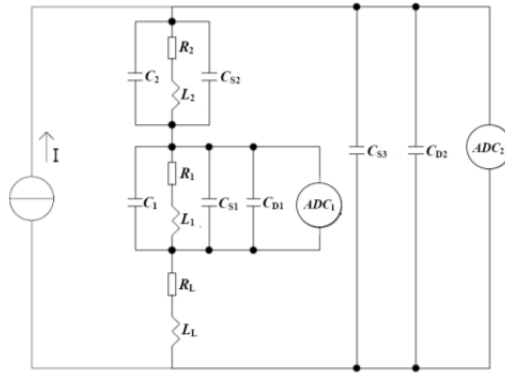


Fig. 7.1.2B. Phase shift measurement step 2.

In Fig.7.1.2A, the shunts, one's phase shift known and another's unknown, are connected in series. Two ADCs simultaneously measure the voltages: one for a shunt near the earth point, and another for the full voltage. The relevant phase difference of  $\varphi_1-\varphi_2$  as the measurement result can be obtained according to Eq. (7.1.2). Fig.7.1.2B shows that the positions of the two shunts are carefully exchanged, and the new phase difference of  $\varphi_3-\varphi_4$  can be obtained.

The influence function of  $C_{D1}$  and  $C_{D2}$  can be eliminated in the difference of both results of  $\Delta\varphi = (\varphi_1-\varphi_2) - (\varphi_3-\varphi_4)$ . Suitable wiring may enable  $C_{S1}$ ,  $C_{S2}$  to be disregarded, and yields

$$\Delta\varphi \approx \omega(L_2 / R_2 - R_2 C_2) - \omega(L_1 / R_1 - R_1 C_1) . \quad (7.1.7)$$

On the basis of the phase shift definition of the shunt in Eq. (7.1.5), we infer that the result from Eq. (7.1.7) requires measurement, and then the phase shift of the unknown shunt is determined using the known shunt.

## 7.2 Phase shift of voltage dividers

### 7.2.1 Phase shift measurement

The phase shift of the voltage divider between its output and input, and relative uncertainty is determined in an experiment. Every one of the sets of dividers at the ranges 8, 15, 30, 60, 120, 240, and 500 V is measured. The experiment is based on a step-up procedure, discussed below.

**Step-up Procedure** Fig. 7.2.1 depicts the step-up procedure for measuring the phase shift of the dividers.

**Preparation:** A voltage of 0.8 V is connected in parallel to both DVMs to eliminate their intrinsic phase difference; that is, the  $0^\circ$  standard is formulated.

**Step 1.** A voltage of 0.8 V is applied parallel to both DVMs (DVM2 at 10 V and DVM1 at 1 V) without any divider. The phase difference between the 10 and 1 V range of DVM2 is measured.

**Step 2.** A voltage of 8 V is applied parallel to the 8 V divider (a divider with a range of 8 V) and DVM2 in its 10 V range. The output of the divider (i.e., 0.8 V) is connected to

DVM1 in its 1 V range. The phase difference between the output and input of the 8 V divider is measured.

**Step 3.** The 15 V divider is measured, and the 8 V divider is taken as a standard. A voltage of 8 V is applied parallel to both dividers, and the outputs of both dividers are connected to two DVMs. The phase shift of the 15 V divider is measured.

Similar procedures are repeated up to the 480 V divider. The measurement results are shown in Table 7.2.1 and their standard deviations are listed in Table 7.2. 2.

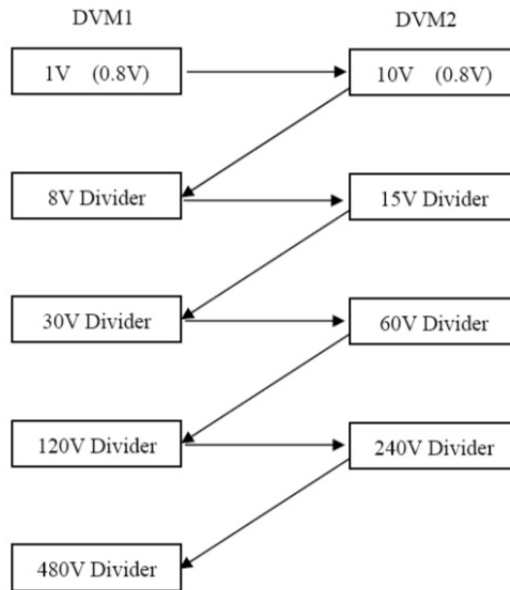


Fig. 7.2.1. Step-up procedure for measuring the phase shift of the voltage dividers.

Order	1	10	20	30	40	50	60
60 V	13.0	131.9	260.3	396.1	515.4	639.5	785.9
120 V	14.6	143.0	282.7	430.2	559.4	694.0	851.3
240 V	12.2	118.2	233.4	356.8	459.8	567.7	701.9
480 V	-14.4	-150.8	-305.7	-451.1	-614.9	-776.7	-918.4

Table 7.2.1. Test results of phase shift  $\varphi_{vk}/\mu\text{rad}$

Order	1	10	20	30	40	50	60
60 V	0.0	0.5	2.7	0.6	7.1	1.1	10.3
120 V	0.3	4.0	5.9	10.3	12.5	14.7	20.9
240 V	0.4	0.5	1.5	1.3	2.4	4.4	3.0
480 V	0.6	1.4	3.7	6.1	8.7	14.2	4.7

Table 7.2.2. Standard deviation  $/\mu\text{rad}$



### 7.2.2 Uncertainty of phase shift measurement

The uncertainty of the phase shift is evaluated according to Eq. (7.1.4).

The  $0^\circ$  standard uncertainty, determined in an experiment, is shown in Table 7.2.3.

Order	1	10	20	30	40	50	60
$u(\Delta\phi)$	3.01	2.12	1.84	2.06	1.70	3.14	3.66

Table 7.2.3. Uncertainty of the  $0^\circ$  standard/ $(\mu\text{rad})$  ( $k = 1$ )

Another factor, the half range effect, is considered. A divider obtains the value in its half range but provides the value in its full range to the next divider.

This effect can be measured by comparing two cases: the full and 50% ranges. The results are shown in Table 7.2.4.

Order	1	10	20	30	40	50	60
$\Delta\phi$	2.0	2.0	3.0	4.0	5.0	6.0	6.5

Table 7.2.4. Variation of phase difference between half and full range/ $(\mu\text{rad})$

Thus, Table 7.2.5 shows the uncertainty of the phase shift of every divider, including the standard deviation in Table 7.2.2, uncertainty of the  $0^\circ$  standard in Table 7.2.3, half range effect in Table 7.2.4, and uncertainty of its foregoing divider.

Order	1	10	20	30	40	50	60
60 V	6.03	5.41	7.64	9.22	11.4	13.9	15.2
120 V	6.38	6.11	8.56	10.6	12.9	15.6	17.2
240 V	6.71	6.47	9.16	11.3	14.0	16.9	18.5
480 V	7.04	6.91	10.3	13.5	17.2	22.8	20.1

Table 7.2.5. Uncertainty of phase shift of dividers after correction,  $u(\Delta\phi_{vk})/(\mu\text{rad})$  ( $k = 1$ )

The 120 V divider under a 100 V fundamental has a phase shift of  $14.6 \mu\text{rad}$ , and its uncertainty is  $13 \mu\text{rad}$  ( $k = 2$ ). However, these results can be applied only to the sinusoidal signal and in full range because the experiments are implemented under these conditions.

### 7.3 Phase shift of shunts

A similar step-up procedure (Fig. 7.3.1) is conducted for the shunts from 0.1 to 20 A, and 50 A. An original standard is necessary; that is, a standard resistor of a known time constant developed early at the NIM. With the parameters  $R = 10 \Omega$ ,  $\tau = 0.03 \times 10^{-9}$ , phase shift is less than  $6 \times 10^{-7}$  within 3 kHz.

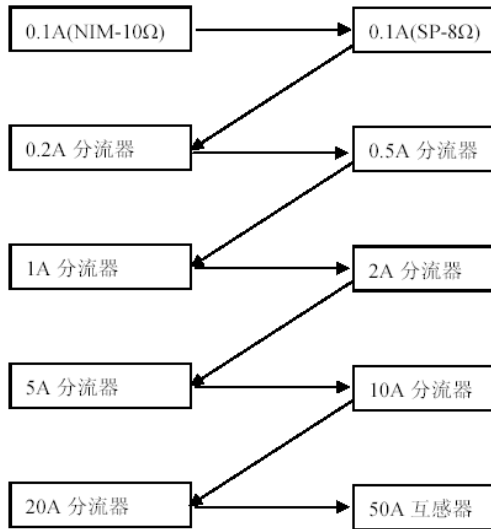


Fig. 7.3.1. Step-up procedure for measuring the phase shift of shunts.

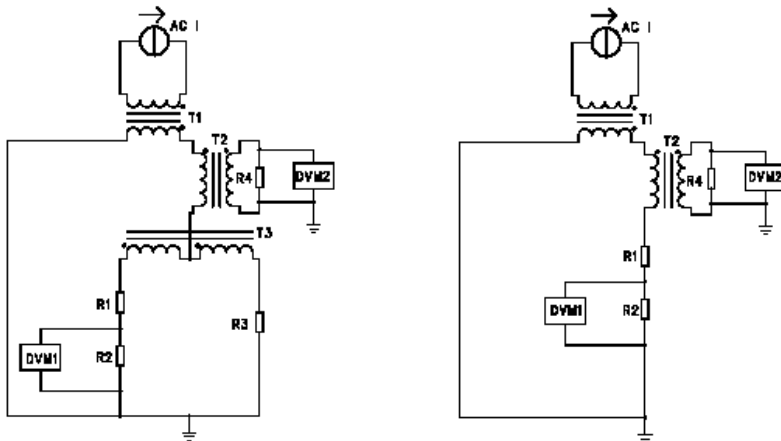


Fig. 7.3.2. Measurement the half range effect of the shunts.

All the shunts and 50 A transformers are measured using the method described in Section 7.1.3. The measurement results for all the shunts are obtained. The results for the shunt of 5 A, including its phase shift and uncertainty, are expressed in Table 7.3.1.

In the uncertainty, all the factors are considered and computed. The half range effect is determined in an experiment, as shown in Fig. 7.3.2.  $R_2$  is measured initially at the half

current of  $0.5I$ , and then at the full current of  $I$ . The half range effect can be determined when  $R_4$  is used as reference.

Order	1	10	20	30	40	50	60
$\Delta\varphi_{lk}$	-1.3	-3.0	-0.7	-.3	-7.0	-12.4	-0.5
$u(\Delta\varphi_{lk})$	5.99	5.38	7.64	10.1	12.4	15.2	16.9

Table 7.3.1. Phase shift and its uncertainty (after correction) of the 5A shunt,  $\mu\text{rad}$ , ( $k = 1$ )

The 5 A shunt under a 5 A fundamental has a phase shift of  $-1.3 \mu\text{rad}$ , and its uncertainty is  $12 \mu\text{rad}$  ( $k = 2$ ). The results can be applied only to the sinusoidal signal and in full range.

## 7.4 Conclusion

The uncertainty evaluation is discussed to measure phase difference by applying DFT to sampling data in NIPS. The step-up procedures are described for the voltage dividers and shunts. The phase shifts and their uncertainties are given. The uncertainty of the phase shift in the fundamental is  $13 \mu\text{rad}$  ( $k = 2$ ) for the 120 V divider, and  $12 \mu\text{rad}$  ( $k = 2$ ) for the 5 A shunt. These specifications fall under the special frequency characteristic. They are discussed further when the results are applied to harmonic power measurement.

## 8. Harmonic power measurement

### 8.1 Determination of Harmonic Power

In the Section 5, DC component  $V_0$ , fundamental component  $V_1$  and its phase angle  $\varphi_{V1}$ , and  $k$ -th harmonic components  $V_k$  and their phase angle  $\varphi_{Vk}$ , ( $k = 2, 3, \dots, w$ ) are obtained for a non-sinusoidal voltage signal.

Section 6 details the derivation of DC component  $I_0$ , fundamental component  $I_1$  and its phase angle  $\varphi_{I1}$ ,  $k$ -th harmonic components  $I_k$  and their phase angle  $\varphi_{Ik}$ , and ( $k = 2, 3, \dots, w$ ) for a non-sinusoidal current signal.

From these, we can calculate the harmonic power.

DC power is

$$P_0 = V_0 I_0. \quad (8.1.1)$$

Fundamental active power, reactive power, and apparent power are

$$P_1 = V_1 I_1 \cos\varphi_{1}, \quad (8.1.2)$$

$$Q_1 = V_1 I_1 \sin\varphi_{1}, \quad (8.1.3)$$

$$S_1 = V_1 I_1. \quad (8.1.4)$$

The  $k$ -th harmonic active power, reactive power, and apparent power are as follows:

$$P_k = V_k I_k \cos\varphi_k, \quad (8.1.5)$$

$$Q_k = V_k I_k \sin \varphi_k, \quad (8.1.6)$$

$$S_k = V_k I_k, \quad (8.1.7)$$

where  $\varphi_1 = \varphi_{V1} - \varphi_{I1}$ ,  $\varphi_k = \varphi_{V_k} - \varphi_{I_k}$ , and  $k = 2, 3, \dots, w$ .

## 8.2 Uncertainty analysis

According to the definition in Section 4.1, the *fiducial error of harmonic active power* can be written as

$$\delta P_k = \Delta P_k / (V_1 I_1), \quad (8.2.1)$$

where  $k = 1, 2, \dots, w$ .

The cases of reactive power and apparent power can be similarly analyzed.

The uncertainty of  $\delta P_k$ , *fiducial uncertainty of harmonic active power*,  $u(\Delta P_k / (V_1 I_1))$  in reference to fundamental apparent power has four parts:

$$u^2(\Delta P_k / (V_1 I_1)) = u_1^2 + u_2^2 + u_3^2 + u_4^2, \quad (8.2.2)$$

where

$$u_1 = (I_k / I_1) \cos \varphi_k u(\Delta V_k / V_1), \quad (8.2.3)$$

$$u_2 = (V_k / V_1) \cos \varphi_k u(\Delta I_k / I_1), \quad (8.2.4)$$

$$u_3 = [\sin \varphi_k I_k V_k / (V_1 I_1)] u(\Delta \varphi_{V_k}), \quad (8.2.5)$$

$$u_4 = [\sin \varphi_k I_k V_k / (V_1 I_1)] u(\Delta \varphi_{I_k}). \quad (8.2.6)$$

Eqs. (8.2.3) and (8.2.4) are dependent on the fiducial uncertainty of the harmonic voltage and current, or  $u(\Delta V_k / V_1)$  and  $u(\Delta I_k / I_1)$ , respectively. They result from the ratio errors of the dividers and shunts.

Eqs. (8.2.5) and (8.2.6) are dependent on the relative uncertainty of the phase difference of the harmonic voltage and current, or  $u(\Delta \varphi_{V_k})$  and  $u(\Delta \varphi_{I_k})$ , respectively. These are analyzed in Section 7. They result from phase angle errors.

When  $\varphi_k = 0^\circ$ , Eqs. (8.2.5) and (8.2.6) equal zero; when  $\varphi_k = 90^\circ$ , Eqs. (8.2.3) and (8.2.4) equal zero.

Eq. (8.2.5),  $u_3$ , can be re-written as

$$u_3 = [\sin \varphi_k \times (I_k / I_1)] [u(\Delta \varphi_{V_k}) \times (V_k / V_1)].$$

where  $u(\Delta \varphi_{V_k})$  is a relative uncertainty of the phase difference of harmonic voltage, which is related to  $V_k$ , and its values in the standard equipment are determined by sampling measurement. The product of  $u(\Delta \varphi_{V_k}) \times (V_k / V_1)$  indicates that the reference of  $u(\Delta \varphi_{V_k})$  should be converted from  $V_k$  to  $V_1$ . As a concept of fiducial uncertainty, this product can be called *fiducial uncertainty of the phase difference of voltage*, denoted as  $u(\Delta \varphi_{V_k} / V_1)$ . Similar to the

case of voltage (current), we consider the effects of the small component, leakage, noise, and non-linearity.

$$u(\Delta\phi_{V_k}/V_1) = (V_k/V_1) u(\Delta\phi_{V_k}).$$

Considering the non-linearity of amplitude under the harmonic conditions in Section 5.2, the ratio of  $(V_k/V_1)$  can be converted to  $[0.8(V_k/V_1) + 0.2]$ . With the addition of the resolution component, a new expression is derived:

$$u^2(\Delta\phi_{V_k}/V_1) = [0.8(V_k/V_1) + 0.2]^2 u^2(\Delta\phi_{V_k}) + 6.85^2.$$

Thus, Eq. (8.2.5) becomes

$$u_3 = [\sin\phi_k (I_k/I_1)] u(\Delta\phi_{V_k}/V_1). \quad (8.2.7)$$

Eq. (8.2.6),  $u_4$ , can be similarly treated.

$$u_4 = [\sin\phi_k (V_k/V_1)] u(\Delta\phi_{I_k}/I_1), \quad (8.2.8)$$

$$u^2(\Delta\phi_{I_k}/I_1) = [0.7(I_k/I_1) + 0.3]^2 u^2(\Delta\phi_{I_k}) + 7.28^2.$$

The digital estimation of  $u_3$  and  $u_4$  can be provided.

### 8.3 Uncertainty evaluation

The uncertainty of harmonic active power can be calculated according to Eq. (8.1.5), where parameters  $V_k/V_1$ ,  $I_k/I_1$  are taken from the characterizing signal in Table 4.2.1, and uncertainties  $u(\Delta V_k/V_1)$ ,  $u(\Delta I_k/I_1)$ ,  $u(\Delta\phi_{V_k})$ , and  $u(\Delta\phi_{I_k})$  are taken from the experimental results in Tables 5.4.1, 6.3.1, 7.2.5, and 7.3.1, respectively. For the cases  $\phi_k = 0^\circ, 30^\circ, 60^\circ$ , and  $90^\circ$ , the calculated results are shown in Table 8.3.1.

Order	1	10	20	30	40	50	60
$0^\circ$	42	28	4.4	3.0	1.6	0.63	0.70
$30^\circ$	39	26	4.2	2.9	1.5	0.61	0.65
$60^\circ$	32	22	3.8	2.6	1.3	0.55	0.52
$90^\circ$	28	19	3.6	2.4	1.3	0.52	0.44

Table 8.3.1. Fiducial uncertainty of active power at 500 V, 20 A under the characterizing signal condition,  $U(\Delta P_k/(V_1 I_1))/(\mu\text{W}/\text{VA})$  ( $k = 2$ )

In Table 8.3.1 the maximum estimated uncertainty ( $k = 2$ ) is  $42 \mu\text{W}/\text{VA}$ , which is the uncertainty index of the harmonic active power. At 100 V and 5 A, however, the maximum estimated uncertainty ( $k = 2$ ) is  $36 \mu\text{W}/\text{VA}$ .

In practical measurement, in order to obtain the best uncertainty of measurement, for example in the important comparison, the parameters of  $V_k/V_1$ ,  $I_k/I_1$  may use of the measuring values instead of the values of the characteristic signal. When the measuring signal consists of only a few harmonics, the resolution component in eqn.(5.4.1), (6.3.1), (8.2.7), (8.2.8) may take smaller values according to the practical case. In the NIM's standard

equipment for the non-sinusoidal signal including 1 to 2 harmonics only, under the best experimental conditions, the value may take 1.00 but not 6.85 or 7.28.

Some factors of interest, such as noise and jitter effects, are already included in the experimental results. For jitter, it exists practically as an uncertainty factor in the time point of sampling. However the author takes the concept that the sampling was at the precise time points still according to the Section 3.1, but, as equivalence, the signal instantaneous amplitudes (the sampling values) were attached additional errors. As far as the aliasing effect is concerned, no analysis can be carried out beyond the 40<sup>th</sup> harmonic component; the analysis is restricted because the power source of the standard equipment does not provide harmonic component signals with orders higher than the 40<sup>th</sup> in satisfying the specified maximum harmonic components of the 40<sup>th</sup> order, as per the IEC standard [IEC, 2002, 2003].

## 9. Experimental validation

The specifications of the standard equipment analyzed and determined in the above-mentioned sections are validated in a series of experiments.

### 9.1 Orthogonality check

According to the orthogonality of the trigonometric function, power can be produced when, and only when, voltage and current are at the same frequency. Fundamental voltage and harmonic current do not produce power, a constraint that can be applied to validate the effectiveness of the equipment [Arseneau et al., 1995b, 2001].

Four steps are designed in the experiment to verify the performance of the standard equipment of the NIM. A time division-type power converter with an accuracy of 50 ppm is selected as a reference standard.

**Step 1.** The input voltage and current signals are fundamental.

**Step 2.** Voltage is fundamental, and current is a combination of fundamental and a harmonic.

**Step 3.** Voltage is a combination of fundamental and a harmonic, and current is fundamental.

**Step 4.** Return to Step 1.

Two sets of non-sinusoidal signals are selected on the basis of (a) IEC61000-3-2 [IEC, 2001] (fundamental voltage, 100 V; current, 3 A; 5<sup>th</sup> order voltage, 10%; 5<sup>th</sup> order current, 40%), and (b) NRC [Arseneau et al., 1995a, 1997] (voltage, 120 V with  $THD = 9.5\%$ ; current, 5 A with  $THD = 19.5\%$ ).

With the set of IEC signals at  $PF = 1.0$ , the errors are  $-1.4$  and  $-0.7 \mu W/VA$  with the 5<sup>th</sup> order current harmonics and 5<sup>th</sup> order voltage harmonics included, respectively. With the set of NRC signals, the corresponding errors are  $-0.6$  and  $-0.4 \mu W/VA$ , which can be disregarded.

Two other commercial power meters, one operating under the principle of a thermo-converter and the other under the principle of sampling, are also subjected to the same test. Similar results were obtained for these power meters.

## 9.2 Test on the commercial programmable signal generator

The commercial programmable signal generator not only provides sinusoidal fundamental waveform voltage and current signals, but also enables the modification of these waveforms to include harmonic components for the calibrations of harmonic meters. It is specified as a 0.02% device (a standard source).

The performance of the programmable signal generator is also verified using the harmonic power standard equipment. The generator is programmed to produce the set of IEC signals of fundamental 50 Hz 100 V, 3 A, and  $\cos\phi = 1$ , and 5<sup>th</sup> harmonic 10 V, 1.2 A, and  $\cos\phi = 1$ . The measured results are compared with the set values of the signal generator. The fiducial harmonic errors of voltage, current, and power from the DC to the 7<sup>th</sup> order harmonics are shown in Table 9.2.1 They basically confirm the specifications of the signal generator.

However, a small non-linearity effect in the current output is observed. Larger errors are generated at the DC, 2nd, 4th, and 6th order harmonic currents. This result may be indicative of a non-linearity effect. Given the square of the fundamental component that produces the DC and 2nd order harmonic components, the cross term of the fundamental and 5th order harmonic component produces the 4th and 6th harmonic components. A similar but less noticeable phenomenon can also be observed in the voltage output. The occurrence of the non-linearity effect should be confirmed through other tests.

Order	Voltage/( $\mu\text{V}/\text{V}$ )	Current/( $\mu\text{A}/\text{A}$ )	Power/( $\mu\text{W}/\text{VA}$ )
0	79.5	<b>-237.0</b>	0.0
1	-12.9	-6.6	-19.5
2	12.1	<b>26.3</b>	0.0
3	8.7	7.1	0.0
4	4.9	<b>32.7</b>	0.0
5	-9.8	-6.3	-4.6
6	6.2	<b>51.3</b>	0.0
7	11.0	6.9	0.0

Table 9.2.1. Fiducial harmonic errors of the signal generator at the IEC signal

## 9.3 Comparison with the national energy standard

Although the harmonic power standard equipment is designed for harmonic power measurements, it can also be used to measure power/energy under sinusoidal waveform conditions. A comparison against the National Primary Power Standard of the NIM is conducted under test conditions of 100 V, 5 A, 50 Hz,  $\cos\phi = 1$ , 0.5 lag, 0.5 lead, zero lag, and zero lead. The errors of the harmonic power standard equipment are no more than  $-18.6 \mu\text{W}/\text{VA}$  at all the test conditions. This result is in agreement with previous findings within the evaluated uncertainties of the calibration/experiment.

Bilateral comparisons with the power standards of the National Metrology Institute of Germany PTB and USA NIST are also carried out using a traveling transfer standard under the test conditions 120 V, 5 A, 50 Hz, power factors of unity, 0.5 lag and lead, and zero lag and lead. The agreement is within  $4 \mu\text{W}/\text{VA}$  at all the test points.

## 10. Conclusion

The harmonic power standard equipment of the NIM is introduced. It is based on the digital sampling technique, which does not require synchronous sampling. It features a special algorithm that compensates for the leakage effect caused by asynchronous sampling. When applied to power measurements with harmonic components of up to the 60<sup>th</sup> order, computation time is less than 2 seconds even without a large computer memory.

We propose a new concept of uncertainty expression that is related to the fundamental, and define it as the fiducial harmonic uncertainty. This concept is designed to evaluate harmonic measurements. A characterizing waveform signal is reported for the universality of the harmonic uncertainty claim of the NIM standard.

The harmonic power standard operates at fundamental frequencies ranging from 45 to 65 Hz with harmonic components of up to the 60<sup>th</sup> order, voltage range from 60 to 500 V, current range from 0.1 to 50 A, and any power factor from zero lag through unity to zero lead. The evaluated uncertainties ( $k = 2$ ) of the harmonic voltages relative to the fundamental voltage, harmonic currents relative to the fundamental current, and harmonic power to the fundamental apparent power are 30  $\mu\text{V}/\text{V}$ , 36  $\mu\text{A}/\text{A}$ , and 42  $\mu\text{VA}/\text{W}$ , respectively. These parameters may indicate the function of the principles and the methods described in this chapter, but will not be a limit for the future work.

## 11. Acknowledgments

The author would like to thank Prof. Tang Tongyi (Ph.D. supervisor of the author) of Tsinghua University for his guidance in the investigation of the algorithm, as well as Profs. Zhang Zhonghua and Zhang Deshi of the NIM and Prof. Yang Fusheng of Tsinghua University for their guidance. Appreciation also goes to Mr. Wang Lei of the NIM for his work in the construction of the equipment, as well as Ms. Li Min, Ms. Liu Lijuan, and Mr. Zhou Hao for their contributions to the project. Special thanks to Dr. Eddy So of the National Research Council of Canada for his valuable advice.

## 12. References

- Arseneau R., Filipski P. S., and Zelle J. J., (1995a). A portable and stable source of ac voltage current and power. *IEEE Trans. Instrum. Meas.*, Vol.44, (Apr. 1995). pp.433–435, ISSN 0018-9456
- Arseneau R. and Filipski P. S., (1995b). A calibration system for evaluation the performance of harmonic power analyzer. *IEEE Trans. Power Delivery*, Vol.10, (July 1995). pp.1177–1182
- Arseneau R., Filipski P.S., and Zelle J.J., (1997). An improved three phase digital recorder system for calibrating power instrumentation. *IEEE Trans. Instrum. Meas.*, Nol. 46, (Apr. 1997). pp.399–402, ISSN 0018-9456
- Arseneau R., Sutherland M. E., Zelle J. J., and Svensson S., (2001). Comparison of Non-sinusoidal Calibration Systems at NRC Canada and SP Sweden. *IEEE Trans. Instrum. Meas.*, Vol.50, (Apr. 2001). pp.275–277, ISSN 0018-9456
- BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, (1993a). *International Vocabulary of Basic and General Terms in Metrology*(Second edition), International Organization for Standardization, ISBN92-67-01075-1, Switzerland
- BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, (1993b). *Guide to the Expression of Uncertainty in Measurement*, ISBN:92-67-10188-9, Switzerland
- Dai X.Z. and Gretschek R., (1994). Quasi-Synchronous Sampling Algorithm and Its Applications, *IEEE Transactions on Instru. & Meas.*, Vol.43, No.2, (April 1994). pp.204-209, ISSN 0018-9456



- IEC, International Standard (2001). IEC61000-3-2, Electromagnetic compatibility (EMC)- Part 3-2: Limits- Limits for harmonic current emissions, Switzerland
- IEC, International Standard (2002). IEC61000-4-7, Electromagnetic compatibility (EMC)- Part 4-7: Testing and measurement techniques-General guide on harmonics and interharmonics measurements and instrumentation, for power supply systems and equipment connected thereto, Switzerland
- IEC, International Standard (2003). 62053-21 · Electricity metering equipment (a.c.)- Particular requirements, Part 21: Static meters for active energy, Switzerland
- Ihlenfeld W.K. et al.,(2003). Evaluation of the synchronous generation and sampling technique. *IEEE Trans. Instrum. Meas.*, Vol., 52, No.2, (2003). pp. 371-374, ISSN 0018-9456
- Lu Z.L., (1986). The interpolation and its error analysis for sampling measurement data. *Acta Metrologica Sinica*, Vol. 7, No. 4, (Oct.1986). pp. 314-322, ISSN 0018-94561000-1158
- Lu Z.L., (1988). An Error Estimate for Quasi-Integer-Period Sampling and an Approach for Improving its Accuracy. *IEEE Trans. Instrum. Meas.*, Vol. 37, No. 2, (June 1988). pp.219-222, ISSN 0018-9456
- Lu Z.L., (1991). Study of non-integer-period sampling and development of the standard of power harmonic analyzer, Ph.D. dissertation, Tsinghua Univ., 1991. Available from [http://opac.nlc.gov.cn/F/TISYTLKRPTC65TF5JJNT1GG16PLJ6DUXR8CBF5IEKPYPQ9VPSR-13390?func=full-set-set&set\\_number=301081&set\\_entry=000001&format=999](http://opac.nlc.gov.cn/F/TISYTLKRPTC65TF5JJNT1GG16PLJ6DUXR8CBF5IEKPYPQ9VPSR-13390?func=full-set-set&set_number=301081&set_entry=000001&format=999)
- Lu Z.L., Wang L., and Li M., (2006). Applying DFT of sampling data for determination of phase angle, *2006CPEM Digest*, pp.562-563, ISBN:88-7992-228-9, Torino, July 2006
- Lu Z.L., Wang L., Li M., Liu L.J., and Zhou H., (2008a). A proposal of definition for uncertainty of harmonic voltage and its experimental determination. *2008CPEM Digest*, pp. 446-447, ISBN:978-1-4244-2399-6, Broomfield, CO, Jun. 2008
- Lu Z.L., Wang L., Li M., Liu L.J., and Zhou H., (2008b). Uncertainty estimation of phase shift of voltage dividers and shunts measured by sampling approach. *2008CPEM Digest*, pp.444-445, ISBN:978-1-4244-2399-6, Broomfield, CO, Jun. 2008
- Lu Z.L., Wang L., Li M., Liu L.J., and Zhou H., (2008c). Precision measurement of harmonic current and power at industrial frequency. *2008CPEM Digest*, pp.204-205, ISBN:978-1-4244-2399-6, Broomfield, CO, Jun. 2008
- Lu Z.L., (2008d). Sampling Strategy for Measurement of Periodic Signals. *Electrical Measurement & Instrumentation*, Vol.45, No.514, (Oct.2008). pp.1-6, ISSN 1001-1390
- Lu Z.L., Wang L., Li M., Liu L.J., and Zhou H., (2010). Harmonic power standard at NIM and its compensation algorithm. *IEEE Trans. Instrum. Meas.*, Vol. 59, No. 1, (January 2010). pp.180-187, ISSN 0018-9456
- Svensson S., (1998). Verification of calibration system for power quality instruments. *IEEE Trans. Instrum. Meas.*, Vol. 47, (Oct. 1998). pp.1391-1394, ISSN 0018-9456
- Wang L., Lu Z.L., Li M., Liu L.J., Zhou H., (2008). A New Approach to Determine Phase Shift of Shunt Based on the Sampling Measurement and Substitution Method. *CPEM2008 Digest*, pp.202-203, ISBN:978-1-4244-2399-6, Broomfield, CO, Jun. 2008
- Xue J.H, Yu H.B., Bai J.F., Lu Z.L., (2011). A Novel Algorithm for Harmonic Measurement Based on Multiple-Average, *Acta Metrologica Sinica*, Vol. 32, No. 5, (Sep.2011). pp.400-406, ISSN 0018-94561000-1158
- Zhang J.Q., Zhao X.M. and Sun J.W., (1996). An Automatic Compensation Algorithm of Quasi-Synchronous Sampling for Distorted Power Measurement. *IMTC 96 Digest*, pp. 951-955, Brussels, June, 1996
- Zhang J.T. et al., (2007). A Method on AC Current Measurements Using Binary Inductive Shunts, *Euromet EM, Low Frequency Experts Meeting 2007*, MIKES, Finland, 2007

# RF and Microwave Power Sensor Calibration by Direct Comparison Transfer

Yueyan Shan<sup>1</sup> and Xiaohai Cui<sup>2</sup>

<sup>1</sup>National Metrology Centre, A\*STAR

<sup>2</sup>National Institute of Metrology

<sup>1</sup>Singapore

<sup>2</sup>P.R. China

## 1. Introduction

Many instruments can be used to measure radio frequency (RF) and microwave power. The most accurate one is a power sensor with a meter. The accuracy of RF and microwave power measurement depends on the accuracy of power sensor calibration. This chapter provides calibration methods of RF and microwave power sensor with system setup, modeling, equations, and analyses in different representations, traceability and measurement uncertainty evaluations. From the simple direct comparison transfer method, to the different improvements, and then the general analytical models, the methods provided in this chapter are useful for the guided-wave power sensor calibration in frequency range of a few MHz to several hundred GHz.

## 2. Microwave power standard and traceability of power sensor calibration

Although the chapter is focusing on the power sensor calibration, it has to talk about first the primary microwave power standard so that the traceability of power sensor calibration is clearly defined.

According to the definition, primary standard is derived directly in terms of base units of the International System of Units (SI). Now the prevalently accepted primary microwave power standard is the calorimeter or microcalorimeter, which is a substitution type of primary standard based on heat measurement (Brunetti & Vremera, 2003; Clague, 1995; Cui, X. & Crowley, T. P. (2011); Famton, 1990; JCGM 200:2008; Oldfield, 1989).

The primary microwave power standard determines the effective efficiency and calibration factor through DC power substitution to realize the traceability to SI units. Power is measured in terms of heat capacity and rate of temperature rise. When microwave power is applied to a terminating device or load through transmission line, the microwave energy is absorbed and converted to heat energy, causing the load temperature to change. Similarly, when DC power is applied to the same load, the DC energy is converted to heat, causing the load temperature to change. When the temperature change caused by the DC power is equivalent to that caused by the microwave power, the DC power can be used to precisely determine the corresponding microwave power. This is the principle of DC power substitution. The substitution technique obviates the need for detailed knowledge of heat losses and thermal capacities.

The terminating device may not react in the same way for microwave and DC power absorption, so the effective efficiency  $\eta_e$  is used to perform the correction. In equation (1), microwave power  $P_{HF}$  absorbed by the terminating device is calculated by dividing the substituted DC power  $P_{DC}$  by the effective efficiency  $\eta_e$ :

$$\eta_e = \frac{P_{DC}}{P_{HF}} \quad (1)$$

Since the effective efficiency is independent of the mismatch correction, the calibration factor  $K$  is used to describe both the effective efficiency  $\eta_e$  and mismatch  $\Gamma$  as follows:

$$K = \eta_e \times (1 - |\Gamma|^2) \quad (2)$$

The calibration factor  $K$  is generally used at the time of calibration transfer from a reference standard to an unknown microwave power sensor. It is the focus in the following sections.

The measurement uncertainty is a non-negative parameter characterizing the dispersion of the effective efficiency  $\eta_e$  or the calibration factor  $K$  being attributed to the standards. The uncertainty is evaluated using “law of propagation of uncertainty” following “Guide to the expression of Uncertainty in Measurement” (GUM) (JCGM 100:2008). Evaluation of measurement and calibration uncertainty by Monte Carlo Method (MCM) is to use Monte Carlo simulation in the uncertainty evaluation of output quantities based on “uncertainty probability distribution propagation” (JCGM 101:2008). The following sections will cover both methods for the measurement uncertainty evaluations.

The value of a primary standard is disseminated to a secondary standard through calibration or comparison. Then the reference standard and working power sensors will be calibrated for use. The measurement results through such relations as unbroken chain of calibrations establish the metrological traceability, each contributing to the measurement uncertainty. The traceability is illustrated in Fig. 1. Here for reference purpose we deliberately provide not only the hierarchy, but also the uncertainties typically related. The real uncertainties depend on the frequency band and each laboratory conditions.

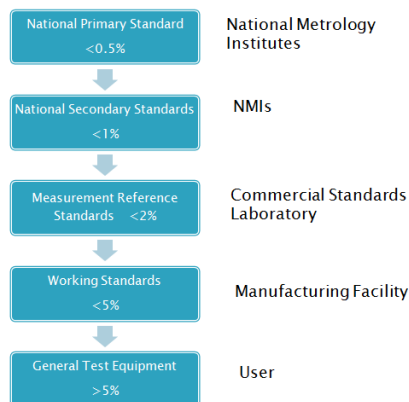


Fig. 1. Dissemination of primary standard to end user- Traceability Chart

### 3. Power sensor calibration by direct comparison transfer

#### 3.1 Modeling

The calibration of RF and microwave power sensor is to transfer the effective efficiency or calibration factor from a primary standard to a secondary standard or from a secondary standard to a reference standard or from a reference standard to a power sensor. The parameter transfer is through comparison, or calibration one against the other. The simplest and the most obvious method to calibrate a power sensor against a reference standard is to connect each in turn to a stable power source, as illustrated in Fig. 2.

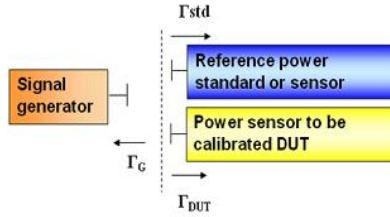


Fig. 2. Calibration of power sensor by the method of simple direct comparison transfer

Generally power from a source,  $P_i$  with reflection coefficient  $\Gamma_G$ , incident to a load with reflection coefficient  $\Gamma_L$ , can be expressed as follows (Agilent, 2003; Engen, 1993; Mial, 2007):

$$P_i = P_{Z_0} \times \frac{1}{|1 - \Gamma_G \Gamma_L|^2}$$

and a reflected power  $P_r$

$$P_r = P_{Z_0} \times \frac{\Gamma_L^2}{|1 - \Gamma_G \Gamma_L|^2}$$

For Fig. 2, the power dissipated to the reference power standard  $P_{Std}$ , can be derived as

$$P_{Std} = P_i - P_{r,Std} = P_{Z_0} \times \frac{1 - |\Gamma_{Std}|^2}{|1 - \Gamma_G \Gamma_{Std}|^2}$$

And the power dissipated to the power sensor to be calibrated (DUT),  $P_{DUT}$ , is as

$$P_{DUT} = P_i - P_{r,DUT} = P_{Z_0} \times \frac{1 - |\Gamma_{DUT}|^2}{|1 - \Gamma_G \Gamma_{DUT}|^2}$$

where  $P_{Z_0}$  is the power available to a load with characteristic impedance  $Z_0$ . When idealized source with available power  $P_a$  has no internal impedance  $Z_0$ ,  $P_{Z_0} = P_a (1 - |\Gamma_G|^2)$ .

As shown in equation (1), the effective efficiency of a power sensor is a ratio.  $P_{DC}$  is DC or low frequency equivalent power, generating the same heat effect as the high frequency power being measured. For calibration of a power sensor, that is, to transfer the effective efficiency of a reference standard,  $\eta_{Std}$ , to a power sensor to be calibrated (DUT),  $\eta_{DUT}$ , it can be derived:

$$\eta_{DUT} = \frac{\frac{P_{DC,DUT}}{P_{RF,DUT}}}{\frac{P_{DC,Std}}{P_{RF,Std}}} = \frac{\frac{P_{DC,DUT}}{P_{Z_0}} \frac{1-|\Gamma_{DUT}|^2}{|1-\Gamma_G\Gamma_{DUT}|^2}}{\frac{P_{DC,Std}}{P_{Z_0}} \frac{1-|\Gamma_{std}|^2}{|1-\Gamma_G\Gamma_{std}|^2}} = \frac{P_{DC,DUT}}{P_{DC,Std}} \times \frac{1-|\Gamma_{std}|^2}{1-|\Gamma_{DUT}|^2} \frac{|1-\Gamma_G\Gamma_{DUT}|^2}{|1-\Gamma_G\Gamma_{std}|^2}$$

So the calibration equation of effective efficiency for the method of Fig. 2 is expressed as:

$$\eta_{DUT} = \eta_{Std} \times \frac{P_{DC,DUT}}{P_{DC,Std}} \times \frac{1-|\Gamma_{std}|^2}{1-|\Gamma_{DUT}|^2} \frac{|1-\Gamma_G\Gamma_{DUT}|^2}{|1-\Gamma_G\Gamma_{std}|^2} \quad (3)$$

where  $\eta_{DUT}$  is the effective efficiency of DUT sensor.

This calibration equation involves three factors:

$$\eta_{DUT} = \eta_{Std} \times P_{RATIO} \times M_1 M_2$$

$\eta_{Std}$  is the effective efficiency of a standard sensor. It comes from a national metrology institute, a calibration laboratory, or a manufacturer with traceability.  $P_{RATIO}$  is an equivalent DC or low frequency power ratio, depending on the system setup.  $M_1 M_2$  is the mismatch factor.

Similarly, calibration equation of calibration factor for the method of Fig. 2 is expressed as follows, considering equation (2):

$$K_{DUT} = K_{Std} \times \frac{P_{DC,DUT}}{P_{DC,Std}} \times \frac{|1-\Gamma_G\Gamma_{DUT}|^2}{|1-\Gamma_G\Gamma_{std}|^2} \quad (4)$$

And the calibration factor transfer equation can also be expressed as three factors:

$$K_{DUT} = K_{Std} \times P_{RATIO} \times M$$

Notice that  $M$  is the mismatch factor for the calibration factor calibration transfer.

### 3.2 Uncertainty evaluation

From calibration transfer equation of effective efficiency (3) and that of calibration factor (4), if DUT power sensor has identical reflection coefficient as that of reference power standard, which means  $\Gamma_{DUT} = \Gamma_{std}$ , each of them absorbs exactly the same amount of power from the source. Then by power ratio measurement it transfers the effective efficiency,  $\eta_e$ , or calibration factor,  $K$ , from standard to DUT, to complete the calibration.

But the actual reflection coefficients of power sensors being compared usually differ significantly from one another (refer to Table 1). The signal generator reflection coefficient cannot dismiss easily. The complex items inside (3) and (4), i.e. the mismatch factors, cannot be neglected in accurate power sensor calibration with small uncertainty.

In recognition of the standards adopted internationally, the GUM (JCGM 100:2008) is considered as the basic technique to evaluate the uncertainty of measurement. The method proposed in the GUM is based on the law of propagation of uncertainty which is essentially the first-order Taylor series approximation of calibration equation, such as equation (4) of calibration factor. The partial differentiation of the output estimate with respect to the input estimates, which is termed sensitivity coefficient, is interpreted as a description of how the output estimate varies with changes in the values of the input estimates. The following equation is the definition of combined standard uncertainty  $u_c(y)$  which includes both type A ( $u_a(x)$ ) and type B ( $u_b(x)$ ) uncertainties when the mathematical model is of the form  $y = f(x_1, x_2, \dots)$ :

$$u_c(y) = \sqrt{u_a^2(x) + \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i}\right)^2 u_b^2(x_i)} \tag{5}$$

Considering that the setup in Fig. 2 is the simple direct comparison transfer method, we present the uncertainty evaluation equations in different implementations, aiming at providing laboratories more realistic choices. The realistic magnitudes for reflection coefficients of instruments are used for the calculation and comparison. The specifications of reflection coefficients of the instruments quoted from different products at some frequency points are listed in Table 1.

product	10 GHz			18 GHz			50 GHz			75GHz(wg)	
	X	Y	Z	X	Y	Z	X	Y	Z	X	Y
sensor (Standard)	0.07	0.01	0.10	0.07	0.03	0.19	0.11	0.02	0.13	0.008	0.006
sensor (DUT)	0.11	0.13	0.16	0.11	0.09	0.06	0.11	0.13	0.20	0.029	0.038
signal generator	0.30	0.33	0.30	0.23	0.33	0.23	0.33	0.33	-	0.26	0.26

Table 1. Typical reflection coefficients of power sensors and signal generators

**3.2.1 The simplest evaluation of measurement uncertainty**

The simplest way in the calibration transfer of calibration factor from standard to DUT sensor is to simplify equation (4) to the following equation:

$$K_{DUT} = K_{Std} \times \frac{P_U}{P_S} = Ratio\ Factor \tag{6}$$

In which only the ratio factor is considered and set the mismatch factor  $M$  equal to 1. The real values of reflection coefficients are considered for the uncertainty budget only. In this case, the mismatch factor  $M$  for calibration factor  $K_{DUT}$  is expressed as:

$$M = \frac{M_U}{M_S}$$

where  $M_U$  and  $M_S$  are given by:

$$M_U = |1 - \Gamma_G \Gamma_{DUT}|^2, \quad M_S = |1 - \Gamma_G \Gamma_{Std}|^2$$

The associated uncertainties are calculated with the following equations (Agilent, 2003; Shan et al., 2010a):

$$u(M_X) = \sqrt{2} |\Gamma_G| |\Gamma_X|; \quad X = Std, DUT \tag{7}$$

According to GUM for the law of propagation of uncertainties, the sensitivity coefficients are partial differentiations with respect to the individual variables in equation (6):

$$\begin{aligned} \frac{\partial K_{DUT}}{\partial K_{Std}} &= \frac{P_U}{P_S} = \frac{K_{DUT}}{K_{Std}} \\ \frac{\partial K_{DUT}}{\partial P_U} &= \frac{K_{Std}}{P_S} = \frac{K_{DUT}}{P_U} \\ \frac{\partial K_{DUT}}{\partial P_S} &= -\frac{K_{Std} \times P_U}{P_S^2} = (-1) \frac{K_{DUT}}{P_S} \end{aligned} \tag{8}$$

In which the expression is prepared for relative uncertainty representation since the combined standard uncertainty  $u_c(y)$  can be expressed as an estimated relative combined

variance  $\left(\frac{u_c(y)}{y}\right)^2 = \sum_{i=1}^N P_i^2 \left(\frac{u(x_i)}{x_i}\right)^2$  when the mathematical model is of the form  $Y = cX_1^{P_1} X_2^{P_2} \dots X_N^{P_N}$ .

The uncertainty budget is listed in Table 2 at frequency 18 GHz with type N connector. The calculations of mismatch uncertainties are based on Table 1 best and worst specifications.

It can be seen from Table 2 that the simplest evaluation of measurement uncertainty method is not an accurate method to obtain a small uncertainty. But it is useful for calibration laboratories with simple measurement set up as illustrated in Fig. 2. When this uncertainty value meets the demand, it is acceptable for industrial applications. In some evaluation experiments, it is also a practical method. Note that the value and uncertainty of a reference standard comes from the calibration of national metrology institutes or other calibration laboratories if the laboratory provides the calibration service using the method.

18 GHz		based on best specifications			
Quant.	Estim.	Standard uncertainty	probability distribution	Sensitivity coefficient	Uncertainty contribution
$X_i$	$x_i$	$u(x_i)$		$c_i$	$u_i(y)$
$K_S$	0.9894	0.0012	normal	1.0137	0.0012
$M_S$	1.0000	0.0098	U-shaped	1.0000	0.0098
$M_U$	1.0000	0.0195	U-shaped	1.0000	0.0195
$P_U$	1.0158	0.0018	normal	0.9873	0.0018
$P_S$	1.0021	0.0004	normal	-1.0008	-0.0004
$y=K_U$	1.0029				0.0219

18 GHz		based on worst specifications			
Quant.	Estim.	Standard uncertainty	probability distribution	Sensitivity coefficient	Uncertainty contribution
$X_i$	$x_i$	$u(x_i)$		$c_i$	$u_i(y)$
$K_S$	0.9894	0.0012	normal	1.0137	0.0012
$M_S$	1.0000	0.0887	U-shaped	1.0000	0.0887
$M_U$	1.0000	0.0513	U-shaped	1.0000	0.0513
$P_U$	1.0158	0.0018	normal	0.9873	0.0018
$P_S$	1.0021	0.0004	normal	-1.0008	-0.0004
$y=K_U$	1.0029				0.1025

Table 2. Uncertainty budget at 18 GHz for simplest evaluation of uncertainty for Fig. 2 measurement setup. Uncertainties  $u_i$  are at one standard deviation. Powers are measured in mW.

### 3.2.2 Measurement uncertainty improvement with mismatch correction

To improve the calibration accuracy and uncertainty evaluation, we have to perform a mismatch correction, i.e., the complex reflection coefficients have to be considered in the calibration model. The complex value is either representing in term of magnitude and phase or its real and imaginary components. In the majority of engineering applications, the magnitude and phase representation (Polar coordinates) is generally preferred because this representation bears a direct relationship to physical phenomena affecting the measurement process (Ridler & Salter, 2002). For example, phase relates directly to the electrical path length of a signal, and magnitude relates directly to the signal attenuation. The same cannot be said for the representation of real and imaginary components (Cartesian coordinates). If the scales are used to depict the different representations, the real and imaginary axes in the complex plane extend infinitely ( $\pm\infty$ ). It is the same as the scale is used to depict all real numbers which is routinely for the arithmetic operations. While the scales are used to represent magnitude and phase each possess a significant difference. The magnitude has a lower bound of zero below which values cannot exist, and phase is represented conventionally on a cyclical scale, either from  $-180^\circ$  to  $+180^\circ$  or from  $0^\circ$  to  $360^\circ$ .

With mismatch correction added in computing the calibration factor using equation (4), the calculation of sensitivity coefficients involves the partial differentiations with respect to complex reflection coefficients. In the following sections, we separately provide the derived sensitivity coefficients in both Cartesian and Polar coordinate representations for practical measurement uncertainty solution and application; and also examples for their associated uncertainties are included. The derived sensitivity coefficients in both representations have been numerically appreciated by making use of MATLAB version R2010a, symbolic differentiation in Math Toolbox.

#### 3.2.2.1 Cartesian representation of sensitivity coefficients for equation (4)

Representing calibration factor  $K_{DUT}$  of equation (4) with magnitude and phase components, it becomes

$$K_{DUT} = K_{Std} \times P_{RATIO} \times M = K_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{1 + |\Gamma_{DUT}|^2 |\Gamma_G|^2 - 2 |\Gamma_{DUT}| |\Gamma_G| \cos(\theta_{DUT} + \theta_G)}{1 + |\Gamma_{Std}|^2 |\Gamma_G|^2 - 2 |\Gamma_{Std}| |\Gamma_G| \cos(\theta_{Std} + \theta_G)} \quad (9)$$



Let  $M_N$  represent the numerator and  $M_D$  denominator of mismatch  $M$ :

$$M_N = |1 - \Gamma_{DUT} \Gamma_G|^2 = 1 + |\Gamma_{DUT}|^2 |\Gamma_G|^2 - 2 |\Gamma_{DUT}| |\Gamma_G| \cos(\theta_{DUT} + \theta_G)$$

$$M_D = |1 - \Gamma_{Std} \Gamma_G|^2 = 1 + |\Gamma_{Std}|^2 |\Gamma_G|^2 - 2 |\Gamma_{Std}| |\Gamma_G| \cos(\theta_{Std} + \theta_G)$$

Let  $A = |\Gamma_G|$ ,  $B = |\Gamma_{DUT}|$ ,  $C = |\Gamma_{Std}|$ ,  $D = \cos(\theta_{DUT} + \theta_G)$ ,  $E = \cos(\theta_{Std} + \theta_G)$ , and  
*Ratio Factor* =  $K_{Std} \times \frac{P_{DUT}}{P_{Std}}$ .

Then,

$$\begin{aligned} K_{DUT} &= K_{Std} \times P_{RATIO} \times M \\ &= \text{Ratio Factor} \times M = \text{Ratio Factor} \times \frac{1 + A^2 B^2 - 2ABD}{1 + A^2 C^2 - 2ACE} \end{aligned} \quad (10)$$

According to GUM for the law of propagation of uncertainties, the sensitivity coefficients are partial differentiations with respect to the individual variables in equation (10), total 9 items. The derived sensitivity coefficients are as follows:

1. The sensitivity coefficient for  $K_{Std}$

$$\frac{\partial K_{DUT}}{\partial K_{Std}} = \frac{K_{DUT}}{K_{Std}} \quad (11)$$

2. The sensitivity coefficient for  $P_{DUT}$

$$\frac{\partial K_{DUT}}{\partial P_{DUT}} = \frac{K_{DUT}}{P_{DUT}} \quad (12)$$

3. The sensitivity coefficient for  $P_{Std}$

$$\frac{\partial K_{DUT}}{\partial P_{Std}} = -\frac{K_{DUT}}{P_{Std}} \quad (13)$$

4. The sensitivity coefficient for  $|\Gamma_{Std}|$  (=C)

$$\frac{\partial K_{DUT}}{\partial |\Gamma_{Std}|} = -\frac{K_{DUT}}{M_D} \times 2A(AC - E) \quad (14)$$

5. The sensitivity coefficient for  $\theta_{Std}$

$$\frac{\partial K_{DUT}}{\partial \theta_{Std}} = -\frac{K_{DUT}}{M_D} \times 2AC \sin(\theta_{Std} + \theta_G) \quad (15)$$

6. The sensitivity coefficient for  $|\Gamma_{DUT}|$  (=B)

$$\frac{\partial K_{DUT}}{\partial |\Gamma_{DUT}|} = \text{Ratio Factor} \times \frac{1}{M_D} \times 2A(AB - D) \quad (16)$$

7. The sensitivity coefficient for  $\theta_{DUT}$

$$\frac{\partial K_{DUT}}{\partial \theta_{DUT}} = \text{Ratio Factor} \times \frac{2AB}{M_D} \sin(\theta_{DUT} + \theta_G) \quad (17)$$

8. The sensitivity coefficient for  $|\Gamma_G| (=A)$

$$\frac{\partial K_{DUT}}{\partial |\Gamma_G|} = \text{Ratio Factor} \times \frac{2B}{M_D} \times (AB - D) - \frac{M_N \times 2C}{(M_D)^2} \times (AC - E) \quad (18)$$

9. The sensitivity coefficient for  $\theta_G$

$$\begin{aligned} \frac{\partial K_{DUT}}{\partial \theta_{DUT}} &= \text{Ratio Factor} \times \left\{ \frac{2AB}{M_D} \sin(\theta_{DUT} + \theta_G) - \frac{M}{M_D} \times 2AC \sin(\theta_{Std} + \theta_G) \right\} \\ &= \text{equation}(17) + \text{equation}(15) \end{aligned} \quad (19)$$

### 3.2.2.2 Polar representation of sensitivity coefficients for equation (4)

Representing calibration factor  $K_{DUT}$  of equation (4) with real and imaginary components, the following denotation is used to denote the real and imaginary components of reflection coefficient:

$$\begin{aligned} a &= \text{Re}\{\Gamma_{Std}\} = \Gamma_{Std-Re}, \quad b = \text{Im}\{\Gamma_{Std}\} = \Gamma_{Std-Img}, \quad c = \text{Re}\{\Gamma_{DUT}\} = \Gamma_{DUT-Re}, \quad d = \text{Im}\{\Gamma_{DUT}\} = \Gamma_{DUT-Img}, \\ e &= \text{Re}\{\Gamma_G\} = \Gamma_{G-Re}, \quad \text{and } f = \text{Im}\{\Gamma_G\} = \Gamma_{G-Img} \end{aligned}$$

Then the calibration model becomes:

$$\begin{aligned} K_{DUT} &= K_{Std} \times P_{RATIO} \times M1 \\ &= \text{Ratio Factor} \times \frac{1+2df-2ce+c^2e^2+d^2e^2+c^2f^2+d^2f}{1+2bf-2ae+a^2e^2+b^2e^2+a^2f^2+b^2f^2} \end{aligned} \quad (20)$$

Let  $M1_N$  represent the numerator and  $M1_D$  denominator of mismatch  $M1$  in real and imaginary format:

$$\begin{aligned} M1_N &= |1 - \Gamma_{DUT} \Gamma_G|^2 = 1 + 2df - 2ce + c^2e^2 + d^2e^2 + c^2f^2 + d^2f \\ M1_D &= |1 - \Gamma_{Std} \Gamma_G|^2 = 1 + 2bf - 2ae + a^2e^2 + b^2e^2 + a^2f^2 + b^2f^2 \end{aligned}$$

Similarly, according to GUM for the law of propagation of uncertainties, the sensitivity coefficients are partial differentiation with respect to each input quantities in equation (20), total 9 items. The derived sensitivity coefficients are as follows:

1. The sensitivity coefficients for  $K_{Std}$ ,  $P_{DUT}$ , and  $P_{Std}$  are the same as those in equations (11), (12) and (13). But here  $K_{DUT}$  should use equation (20) instead.
2. The sensitivity coefficient for  $\Gamma_{Std-Re} (= a)$  can be evaluated to obtain

$$\frac{\partial K_{DUT}}{\partial \Gamma_{Std-Re}} = \frac{K_{DUT}}{M1_D} \times (2e - 2ae^2 - 2af^2) \quad (21)$$

3. The sensitivity coefficient for  $\Gamma_{Std-Img}$  ( $= b$ )

$$\frac{\partial K_{DUT}}{\partial \Gamma_{Std-Img}} = \frac{K_{DUT}}{M1_D} \times \{(2f + 2be^2 + 2bf^2)\} \quad (22)$$

4. The sensitivity coefficient for  $\Gamma_{DUT-Re}$  ( $= c$ )

$$\frac{\partial K_{DUT}}{\partial \Gamma_{DUT-Re}} = Ratio\ Factor \times \frac{2e + 2ce^2 + 2cf^2}{M1_D} \quad (23)$$

5. The sensitivity coefficient for  $\Gamma_{DUT-Img}$  ( $= d$ )

$$\frac{\partial K_{DUT}}{\partial \Gamma_{DUT-Img}} = Ratio\ Factor \times \frac{2f + 2de^2 + 2df^2}{M1_D} \quad (24)$$

6. The sensitivity coefficient for  $\Gamma_{G-Re}$  ( $= e$ )

$$\frac{\partial K_{DUT}}{\partial \Gamma_{G-Re}} = Ratio\ Factor \times \frac{2c + 2c^2e + 2d^2e + (M1)(2a - 2a^2e - 2b^2e)}{M1_D} \quad (25)$$

7. The sensitivity coefficient for  $\Gamma_{G-Img}$  ( $= f$ )

$$\frac{\partial K_{DUT}}{\partial \Gamma_{G-Img}} = Ratio\ Factor \times \frac{2d + 2c^2f + 2d^2f - (M1)(2b + 2a^2f + 2b^2f)}{M1_D} \quad (26)$$

### 3.2.2.3 Example

With the same data as used in Table 2, the uncertainty budget is listed in Table 3 with mismatch corrections considered in the calibration equation.

18 GHz	based on best specifications				
Quant.	Estim.	Standard uncertainty	probability distribution	Sensitivity coefficient	Uncertainty contribution
$X_i$	$x_i$	$u(x_i)$		$c_i$	$u_i(y)$
$K_{Std}$	0.9894	0.0012	normal	0.9996	0.0012
$P_{DUT}$	1.0158	0.0018	normal	0.9737	0.0018
$P_{Std}$	1.0021	0.0004	normal	-0.9870	-0.0004
$\Gamma_{DUTmag}$	0.0600	0.0120	normal	-0.4613	-0.0055
$\Gamma_{DUTPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$\Gamma_{Stdmag}$	0.0300	0.0060	normal	0.4581	0.0027
$\Gamma_{StdPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$\Gamma_{Gmag}$	0.2300	0.0460	normal	-0.0606	-0.0028
$\Gamma_{GPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$y=K_U$	0.9890				0.0071

18 GHz		based on worst specifications			
Quant.	Estim.	Standard uncertainty	probability distribution	Sensitivity coefficient	Uncertainty contribution
$X_i$	$x_i$	$u(x_i)$		$c_i$	$u_i(y)$
$K_{Std}$	0.9894	0.0012	normal	1.0716	0.0013
$P_{DUT}$	1.0158	0.0018	normal	1.0437	0.0019
$P_{Std}$	1.0021	0.0004	normal	-1.0580	-0.0004
$\Gamma_{DUTmag}$	0.1100	0.0220	normal	-0.7261	-0.0160
$\Gamma_{DUTPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$\Gamma_{Stdmag}$	0.1900	0.0380	normal	0.7466	0.0284
$\Gamma_{StdPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$\Gamma_{Gmag}$	0.3300	0.0660	normal	0.1878	0.0124
$\Gamma_{GPhase}$	3.1416	1.5709	normal	0.0000	0.0000
$y=K_U$	1.0602				0.0349

Table 3. Uncertainty budget at 18 GHz for measurement uncertainty improvement with mismatch correction for Fig. 2 measurement setup. Uncertainties  $u_i$  are at one standard deviation. Powers are measured in mW.

Compare the result in Table 3 with that in Table 2, the uncertainty is improved by mismatch correction. In the calculation, the magnitude uncertainty of reflection coefficient is assumed to be 40% of its value and the phase uncertainty is assumed to be  $180^\circ$  and value is  $\pi$  for all. In terms of computation cost, it is the same for both Cartesian and Polar coordinate representations. Note that additional uncertainties should also be included such as the connector repeatability, noise, cable flexibility, drift, linearity and frequency error when they are not negligible in practical application.

### 3.3 Discussions on simple direct comparison transfer

#### 3.3.1 The pros and cons for simple direct comparison transfer

The above analyses based on the method illustrated in Fig. 2 are the simple direct comparison transfer calibration method. It is the most basic microwave power transfer technique. The system setup is fast, easy and simple.

But it is only for relatively rough measurement, since the reflection coefficient of a generator  $\Gamma_G$ , is more difficult to measure than that of a passive load.  $\Gamma_G$  varies with time and frequency. Many signal generators and amplifiers have non-linear output impedance. There are several developments on the measurement methods for  $\Gamma_G$  (Shimaoka et al., 2006; Torok et al., 2001) which is beyond the discussion of this chapter.

#### 3.3.2 Improvement of direct comparison transfer by inserting passive components

To obviate the  $\Gamma_G$ , an intermediate component is proposed to use to increase the accuracy. If put an attenuator or an isolator on the source output, the reflection coefficient of a generator is improved. But low power is a problem for inserting an attenuator and being limited in frequency range is a problem for isolators (Rumfelt & Elwell, 1967).

A leveling circuit or ratio measurement can avoid the troubles. Resistive power splitter or directional coupler is normally used as the intermediate component and constructing the leveling circuit.

### 3.3.3 Splitter vs divider vs Tee: Scattering parameter matrix and suitable applications

Before going further analyses, it is time to say a few words about the power splitter (Johnson, 1975), power divider and Tee since even some experienced engineers misuse them and not clear of the differences among them fundamentally. The physical and mathematical expressions are given here.

As illustrated in Fig. 3 for a) power divider; b) power splitter; and c) Tee, there are three resistors for power divider; two resistors for power splitter; and no resistors for Tee.

The three resistors in power divider are  $Z_0/3$  each; for 50 ohm system, they are  $16\text{-}2/3\ \Omega$ . Two resistors in power splitter are  $Z_0$  each.

The power splitter has fixed input at port one; power divider and Tee inputs are exchangeable (bi-directional).

Substantially the scattering parameter matrix or S-parameter matrix are different for these three 3-port components as shown under the physical structures in Fig. 3; in which the S-parameter matrixes are for ideal cases.

Power splitter is used in leveling or ratio measurement; power divider is used for simple power division. Tee is used in low frequency case and in those non-critical measurements.

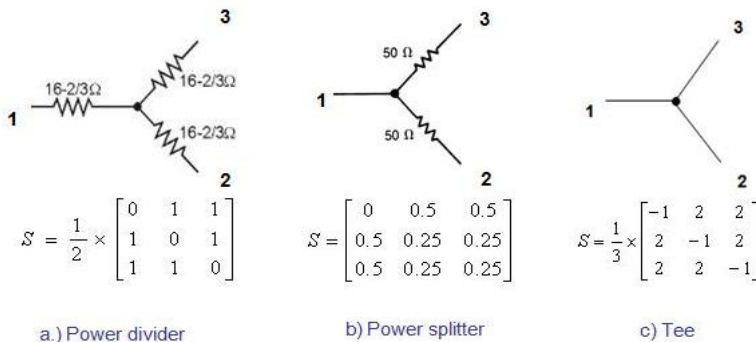


Fig. 3. Three 3-port components and their corresponding scattering parameter matrix

## 4. Coaxial splitter based power sensor calibration by direct comparison transfer

### 4.1 Benefit

The coaxial splitter based calibration setup is illustrated in Fig. 4. By using coaxial splitter in the direct comparison transfer method of RF and microwave power sensor calibration, the measurement accuracy improves from source mismatch effect and the load / device under test (DUT) mismatch effects. The explanations are as follows:

- a. Improving source match is achieved by holding effective source output power constant. When a leveling loop or ratio sensor is employed, port 2 of the power splitter becomes the effective source output.  $\Gamma_{EG}$  is the equivalent source reflection coefficient rather than  $\Gamma_G$ , as illustrated in Fig. 5(a). It solves the measurement trouble of  $\Gamma_G$  for signal generator by obtaining passive component splitter S-parameters. The well known equivalent source reflection coefficient equation for power splitter

$$\Gamma_{EG} = S_{22} - S_{21} \frac{S_{32}}{S_{31}} \tag{27}$$

will be further discussed in section 6.

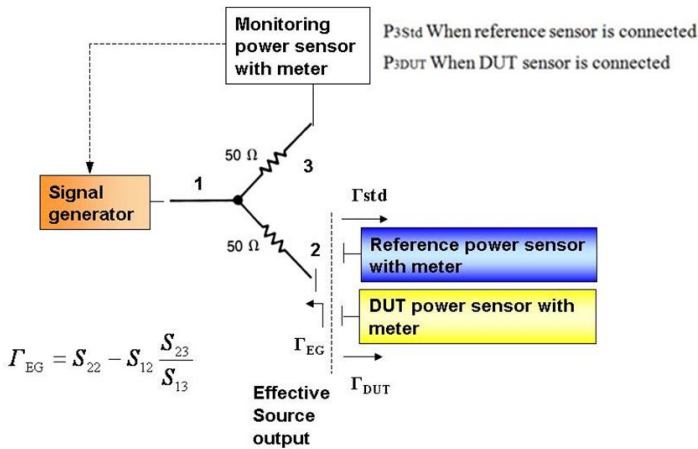


Fig. 4. Coaxial splitter based power measurement setup

- b. As illustrated in Fig. 5(b), since  $S_{21} = S_{31}$ , both arms 2 and 3 experience the same variation in input power, the leveling loop can compensate for any changes in effective source’s output power.
- c. As illustrated in Fig. 5(c) the leveling loop compensates for load variations as well. Since  $S_{22} = S_{32}$  it means that any changes in output power caused by the load reflection is also seen in the monitoring arm, permitting the leveling loop to compensate for these variations.

The source mismatch effect can be read through the monitoring arm, and so do the load / DUT mismatch effects. The system setup and mathematical model is established based on the complex reflection coefficients. Analyses have been provided through their S-parameter matrix to perform full mismatch corrections (Weidman, 1996; Juroshek, 2000; Ginley, 2006; Crowley, 2006; Shan et al., 2008; Shan et al., 2010b). Here we provide models for different cases in the application.

Since the splitter is not perfect, we seek to characterize them so that the mismatch error can be minimized by mathematical correction. The computation involved is not trivial, but the benefits of the technique are considerable, in that accurate measurements are made across a broad band without the need for mechanical adjustment.

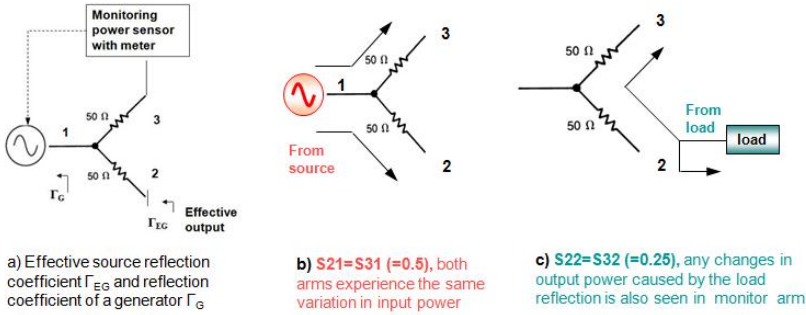


Fig. 5. Benefits by inserting coaxial power splitter

### 4.2 Modeling

Models of the three cases are considered for the system setup illustrated in Fig. 4.

Case 1 - obtain  $\eta_{DUT}$  from  $\eta_{Std}$ :

$$\eta_{DUT} = \eta_{Std} \times P_{RATIO} \times MM = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{1 - |\Gamma_{Std}|^2}{1 - |\Gamma_{DUT}|^2} \times \frac{|1 - \Gamma_{DUT}\Gamma_{EG}|^2}{|1 - \Gamma_{Std}\Gamma_{EG}|^2} \quad (28)$$

The  $\Gamma_{Std}$ ,  $\Gamma_{DUT}$  and  $\Gamma_{EG}$  are the complex value reflection coefficients for the standard, DUT and equivalent signal generator respectively as indicated in Fig. 4. The power ratio is different from equation (3), which is caused by using splitter with monitoring arm. And here  $\Gamma_{EG}$  is used instead of  $\Gamma_G$ , different from simple direct comparison transfer. The value of  $\Gamma_{EG}$  is obtained by equation (27).

Case 2 - obtain  $K_{DUT}$  from  $K_{Std}$ :

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{|1 - \Gamma_{DUT}\Gamma_{EG}|^2}{|1 - \Gamma_{Std}\Gamma_{EG}|^2} \quad (29)$$

Case 3 - obtain  $K_{DUT}$  from  $\eta_{Std}$ :

$$K_{DUT} = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times (1 - |\Gamma_{Std}|^2) \times \frac{|1 - \Gamma_{DUT}\Gamma_{EG}|^2}{|1 - \Gamma_{Std}\Gamma_{EG}|^2} \quad (30)$$

Case 3 is the practical application calibration equation in national metrology institutes. The obtained value from a primary standard is effective efficiency  $\eta_{Std}$ , the customer DUT calibration factor requests  $K_{DUT}$ .

In the following sections, we separately provide both Cartesian and Polar representations for practical solution and application.

### 4.3 Polar model

The models for the above three cases in terms of magnitude and phase is derived and expressed as follows:

$$\eta_{DUT} = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{1 - |\Gamma_{Std}|^2}{1 - |\Gamma_{DUT}|^2} \times \frac{1 + |\Gamma_{DUT}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{DUT}| |\Gamma_{EG}| \cos(\theta_{DUT} + \theta_{EG})}{1 + |\Gamma_{Std}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{Std}| |\Gamma_{EG}| \cos(\theta_{Std} + \theta_{EG})} \quad (31)$$

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{1 + |\Gamma_{DUT}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{DUT}| |\Gamma_{EG}| \cos(\theta_{DUT} + \theta_{EG})}{1 + |\Gamma_{Std}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{Std}| |\Gamma_{EG}| \cos(\theta_{Std} + \theta_{EG})} \quad (32)$$

$$K_{DUT} = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times (1 - |\Gamma_{Std}|^2) \times \frac{1 + |\Gamma_{DUT}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{DUT}| |\Gamma_{EG}| \cos(\theta_{DUT} + \theta_{EG})}{1 + |\Gamma_{Std}|^2 |\Gamma_{EG}|^2 - 2 |\Gamma_{Std}| |\Gamma_{EG}| \cos(\theta_{Std} + \theta_{EG})} \quad (33)$$

#### 4.4 Uncertainty based on polar representation by Monte Carlo method

If using GUM to evaluate the uncertainty of measurement, it is based on propagation of uncertainties which is similar to analyses in previous section with more items. The mismatch uncertainty part is similar, just replace  $\Gamma_G$  with  $\Gamma_{EG}$ . It is seen from previous section that the sensitivity coefficients are quite tedious for complex value involved models although only first order summation of uncertainties are used. The derived sensitivity coefficients with partial differentiations with respect to each variable for the above models (31), (32) and (33) are obtained and have been numerically appreciated by making use of MATLAB. Here we discuss the Monte Carlo simulation Method (MCM). Then we compare the uncertainties by two methods and discuss the findings.

MCM is based on the propagation of distribution proposed by GUM supplement 1 (JCGM, 101:2008) instead of the GUM propagation of uncertainty method. The MCM allows one to get rid of much of the calculation of partial derivatives where analytical expressions are complex. The MCM evaluates measurement uncertainty by setting a probability function to each input quantity in the measurement equation. From a series of numerical calculations, probability density function (pdf) of the output function is obtained and the standard uncertainty is evaluated from this pdf.

The steps for applying MCM are summarized as follows:

- Step 1. Select the number  $N$  of Monte Carlo Trials (iterations) to be made, with a given model  $Y$ , such as using equation (32). A value of  $10^6$  is often expected to deliver a 95% coverage interval for the output quantity.
- Step 2. Generate  $N$  samples, by means of sampling from each input pdf. This  $N$  samples could be stored as a row vector in the program.
- Step 3. For each of the input vector, form the output model  $Y$ , such as using equation (32).
- Step 4. Sort the  $N$  evaluated values into a strict increasing order. Here, strict increasing requires all output values of  $Y$  to be unique. A good random generator in the program should satisfy this condition.
- Step 5. With the sorted values, create a histogram for an approximation of the output pdf.
- Step 6. Evaluate the estimate of the output model, given as  $y$ , and the standard uncertainty of  $y$ , given as  $u(y)$ .
- Step 7. Form a coverage interval for  $Y$ , given a required probability  $p$ . If the output pdf is symmetric, then the probabilistically symmetric interval is equal to the shortest coverage interval. Otherwise, shortest coverage interval has to be used.



The Monte Carlo numerical simulation is performed through program developed using MATLAB software. It allows estimating the measurement uncertainties based on the mathematical models.

#### 4.5 Measurement system setup and type N connection result

The realization and implementation of coaxial splitter based power sensor calibration by direct comparison transfer shown in Fig. 4 is illustrated in Fig. 6 for the type N connector system setup (up to 18 GHz).

The calibration results of calibration factor of a power sensor with the measurement uncertainties are illustrated in Fig. 7. The DUT sensor is calibrated by a power standard (CN mount) in the frequency range 50 MHz to 18 GHz using the direct comparison transfer method with Weinschel 1870A splitter. The effective efficiency of power standard was measured using a micro-calorimeter. The calibration factor is calculated using equation (33). The expanded uncertainties are less than 0.9% for the frequency range at coverage factor of  $k = 2$ .

Compared with the results obtained in section 3, it is shown that with insertion of the passive splitter with monitoring arm, measurement uncertainties are much improved. And such measurement transfer system adds little measurement uncertainty when transfer the value from standard to DUT.

Uncertainty budget table is listed in Table 4 for the calibration system at frequency of 8 GHz.

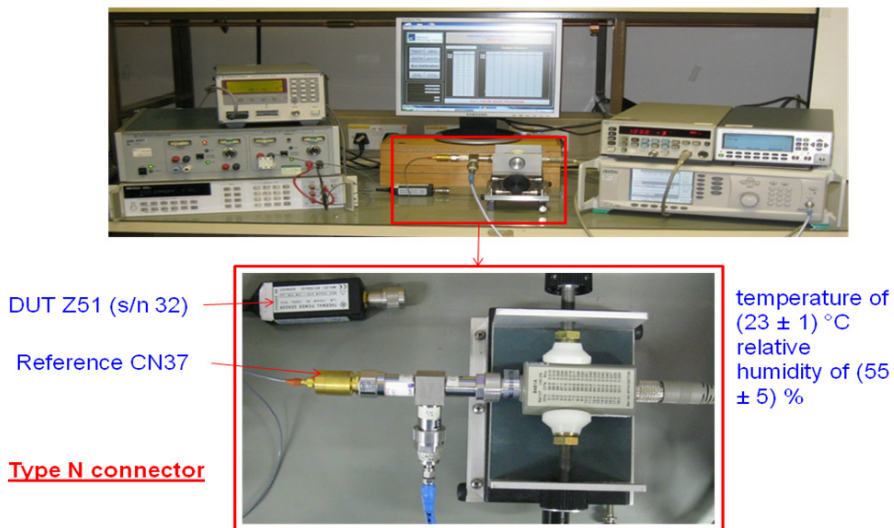


Fig. 6. Type N connection realization of splitter based power sensor calibration by direct comparison transfer

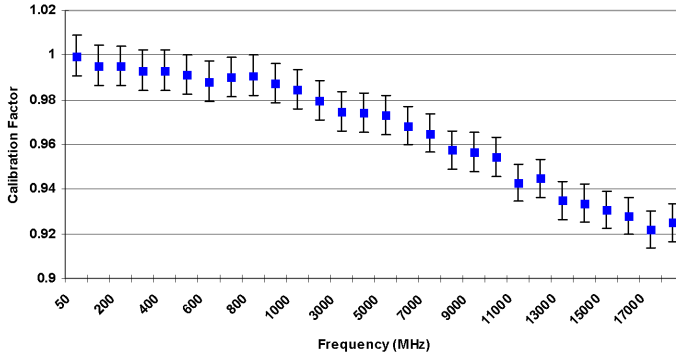


Fig. 7. measurement results by direct comparison method for power sensor 8481A vs CN

Quant.	Estim.	Standard uncertainty	probability distribution	Sensitivity coefficient	Uncertainty contribution
$X_i$	$x_i$	$u(x_i)$		$c_i$	$u_i(y)$
$\eta_{Std}$	0.9650	0.00165	normal	1.00617	0.00166
$P_{Std}$	0.9774	0.00036	normal	-0.99343	-0.00036
$P_{DUT}$	0.9886	0.00171	normal	0.98219	0.00168
$P_{3Std}$	1.0000	0.00010	normal	0.97099	0.00010
$P_{3DUT}$	1.0000	0.00010	normal	-0.97099	-0.00010
$\Gamma_{Stdmag}$	0.0466	0.00750	normal	-0.14611	-0.00110
$\Gamma_{StdPhase}$	-1.4228	0.18328	normal	-0.00269	-0.00049
$\Gamma_{DUTmag}$	0.0047	0.00750	normal	-0.07607	-0.00057
$\Gamma_{DUTPhase}$	2.8563	1.57088	normal	0.00012	0.00020
$\Gamma_{EGmag}$	0.0414	0.00751	normal	-0.07154	-0.00054
$\Gamma_{EGphase}$	-2.5226	0.18381	normal	-0.00257	-0.00047
$y=K_U$	<b>0.9678</b>				<b>0.0029</b>

Table 4. Uncertainty budget at 8 GHz for Fig. 4 measurement setup. Uncertainties  $u_i$  are at one standard deviation. Powers are measured in mW.

4.6 Cartesian model

Real and imaginary expression of the measurement models in the three cases are derived as follows:

$$\eta_{DUT} = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{1 - A^2 - B^2}{1 - C^2 - D^2} \times \frac{1 + 2DF - 2CE + C^2E^2 + D^2E^2 + C^2F^2 + D^2F^2}{1 + 2BF - 2AE + A^2E^2 + B^2E^2 + A^2F^2 + B^2F^2} \quad (34)$$

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times \frac{1 + 2DF - 2CE + C^2E^2 + D^2E^2 + C^2F^2 + D^2F^2}{1 + 2BF - 2AE + A^2E^2 + B^2E^2 + A^2F^2 + B^2F^2} \quad (35)$$

$$K_{DUT} = \eta_{Std} \times \frac{P_{DUT}}{P_{Std}} \times \frac{P_{3Std}}{P_{3DUT}} \times (1 - A^2 - B^2) \times \frac{1 + 2DF - 2CE + C^2E^2 + D^2E^2 + C^2F^2 + D^2F^2}{1 + 2BF - 2AE + A^2E^2 + B^2E^2 + A^2F^2 + B^2F^2} \quad (36)$$

where for all three cases,  $\Gamma_{Std} = A + jB$ ,  $\Gamma_{DUT} = C + jD$ ,  $\Gamma_{EG} = E + jF$ .

**4.7 Measurement system for 2.4mm connection and comparison**

The measurement setup for 2.4mm connector (up to 50 GHz) is illustrated in photo in Fig. 8. The calibration result is illustrated in Fig. 9. For uncertainty evaluation, the parameters used in the MCM are listed in Table 5. And the two columns in Table 6 compare the results obtained from GUM and MCM for selected frequency points.

Using Monte Carlo method, we generate a graphical approximation in the form of a histogram of the probability density function of the output quantity, represented in Fig. 10(a) with the frequency at 50 GHz. The uncertainty differences from the two methods for different frequency range are illustrated in Fig. 10(b). In this case, the uncertainties obtained by different methods are quite close. They can be used as the verification of the measurement uncertainty evaluations.

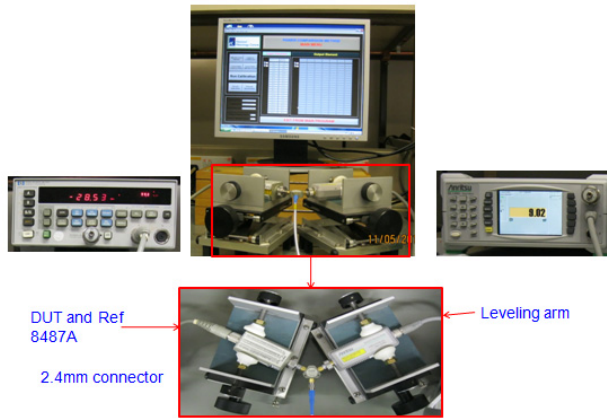


Fig. 8. 2.4mm connection realization of splitter based power sensor calibration by direct comparison transfer

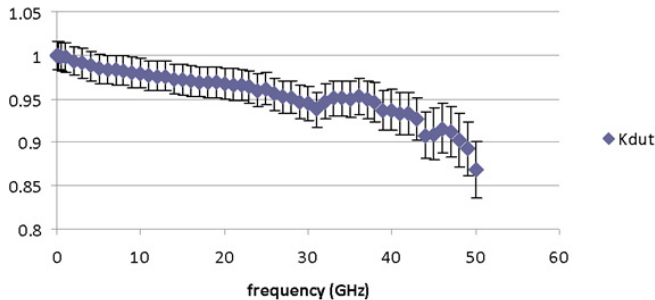


Fig. 9. 50 MHz to 50 GHz using the direct comparison transfer. The expanded uncertainties are less than 4% for the frequency range at  $k=2$ .

$X_i$	Distribution	Parameters	
		Expectation $\mu$	Standard deviation $\sigma$
$\eta_{Std}$	$N(\mu, \sigma^2)$	0.9047	0.0158
$P_{DUT}$	$N(\mu, \sigma^2)$	8.4400e-04	1.0000e-06
$P_{Std}$	$N(\mu, \sigma^2)$	8.6200e-04	1.0000e-06
$P_{3DUT}$	$N(\mu, \sigma^2)$	9.9800e-04	1.0000e-07
$P_{3Std}$	$N(\mu, \sigma^2)$	1.0000e-03	1.0000e-07
$ \Gamma_{DUT} $	$N(\mu, \sigma^2)$	0.1484	0.0125
$\theta_{DUT}$	$N(\mu, \sigma^2)$	2.7501	0.1453
$ \Gamma_{EG} $	$N(\mu, \sigma^2)$	0.1384	0.0125
$\theta_{EG}$	$N(\mu, \sigma^2)$	2.6222	0.1448
$ \Gamma_{Std} $	$N(\mu, \sigma^2)$	0.1288	0.0104
$\theta_{Std}$	$N(\mu, \sigma^2)$	2.9504	0.0630

Table 5. Probability distribution parameters assigned to the individual input quantities at 50 GHz

Frequency = 50MHz		
	GUM	MCM
Mean	1.00854	1.008492
Stdev	0.008106	0.0076298
Lower Lim	0.992328	0.9935218
Upper Lim	1.024753	1.023428
Coverage Int	0.016212	0.0149531

Frequency = 50GHz		
	GUM	MCM
Mean	0.874599	0.8728157
Stdev	0.01635	0.0154997
Lower Lim	0.841899	0.8423992
Upper Lim	0.907299	0.9031398
Coverage Int	0.032699	0.0303703

Table 6. Compares the results of different uncertainty evaluation methods for selected frequency points

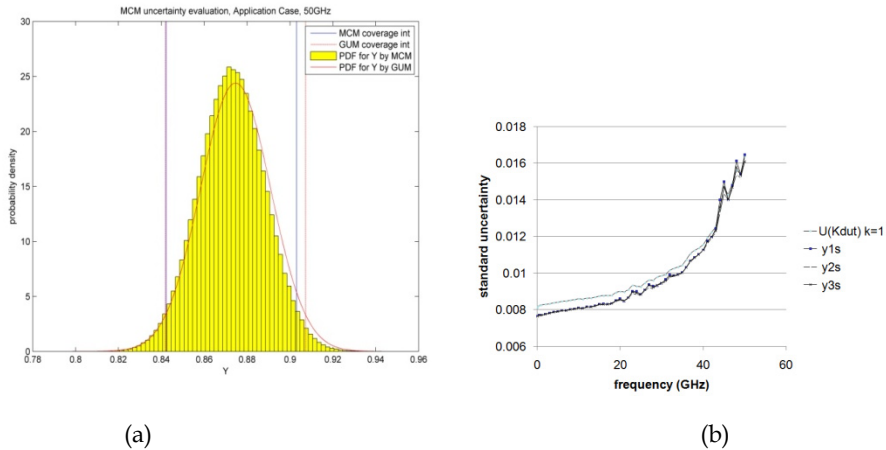


Fig. 10. Comparison of uncertainty analyses by GUM and MCM

### 5. Generic models for power sensor calibration

#### 5.1 Modeling with signal flow graph

For the sake of general purpose of application and analyses, in this section we discuss the three port structure expressed in signal flow graph. A flow graph method is particularly helpful in understanding a complex network relying on S-parameter matrix. Its correspondence with the physical behavior of the circuits allows simplification through well-founded approximations with no loss of physical insight (Bryant, 1993).

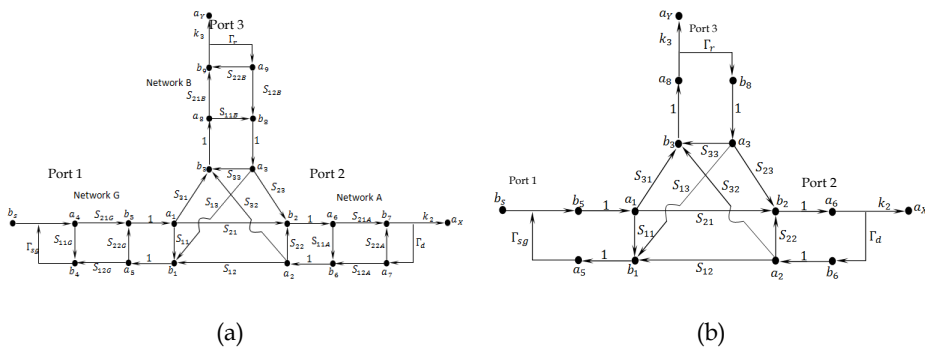


Fig. 11. (a) Generic model signal flow graph for a 3-port structure. (b) Three two-port networks removed from Generic model

Referring to Fig. 11(a), the central part is signal flow graph for a general three port components, most commonly used are splitter and coupler. Three two-port networks, G, A and B, are added for the generic analysis. The signal source connected to port 1 is denoted by  $b_s$ .

## 5.2 Non-touching theory for the modeling analyses

The full power of flow graphs is evident only when they are combined with the non-touching loop rule. It is then possible to find the transfer function between any two nodes of a network almost by inspection. The complex wave functions  $a_i$ ,  $b_i$  are the nodal points of signal flow along the paths designated by the arrows.

There are several special cases which are practically useful. We are going to discuss the following three cases.

### 5.3 Case study 1: Special case of the generic model comparing with section 4

When removing all the three two-port networks, G, A and B, the generic model is the special case as discussed in section 4. The generic model signal flow graph for a 3-port structure is then simplified as Fig. 11(b) (Wong, 2002).

The steps to evaluate the flow graph are summarized as follows:

$$\begin{array}{l}
 \hline
 \text{There are two paths from } b_s \text{ to } a_x \\
 P_1 = S_{21} \cdot k_2 \\
 P_2 = S_{31} \cdot \Gamma_r \cdot S_{23} \cdot k_2 \\
 \hline
 \text{The transfer function is,} \\
 \frac{a_x}{b_s} = \frac{S_{21} \cdot k_2 [1 - (\Gamma_r \cdot S_{33})]}{1 - \sum L(1) + \sum L(2) - \sum L(3) + \dots} + \frac{S_{31} \cdot \Gamma_r \cdot S_{23} \cdot k_2 [1 - 0]}{1 - \sum L(1) + \sum L(2) - \sum L(3) + \dots} \\
 \hline
 \text{There are two paths from } b_s \text{ to } a_y \\
 P_1 = S_{31} \cdot k_3 \\
 P_2 = S_{21} \cdot \Gamma_d \cdot S_{32} \cdot k_3 \\
 \hline
 \text{The transfer function is,} \\
 \frac{a_y}{b_s} = \frac{S_{31} \cdot k_3 [1 - \Gamma_d \cdot S_{22}]}{1 - \sum L(1) + \sum L(2) - \sum L(3) + \dots} + \frac{S_{21} \cdot \Gamma_d \cdot S_{32} \cdot k_3 [1 - 0]}{1 - \sum L(1) + \sum L(2) - \sum L(3) + \dots} \\
 \hline
 \frac{a_x}{a_y} = \frac{a_x}{b_s} \times \frac{b_s}{a_y} \\
 \hline
 \end{array}$$

Taking the ratio of  $a_x$  with respect to  $a_y$ :

$$\begin{aligned}
 \frac{a_x}{a_y} &= \frac{a_x}{b_s} \times \frac{b_s}{a_y} = \frac{k_2}{k_3} \left[ \frac{S_{21}(1 - \Gamma_r \cdot S_{33}) + S_{31} \cdot \Gamma_r \cdot S_{23}}{S_{31}(1 - \Gamma_d \cdot S_{22}) + S_{21} \cdot \Gamma_d \cdot S_{32}} \right] = \left( \frac{k_2}{k_3} \right) \left( \frac{S_{21}}{S_{31}} \right) \left( \frac{1 - \Gamma_r \left( S_{33} - \frac{S_{31} S_{23}}{S_{21}} \right)}{1 - \Gamma_d \left( S_{22} - \frac{S_{21} S_{32}}{S_{31}} \right)} \right) \\
 &= \left( \frac{k_2}{k_3} \right) \left( \frac{S_{21}}{S_{31}} \right) \left( \frac{1 - \Gamma_r \Gamma_{e3}}{1 - \Gamma_d \Gamma_{e2}} \right) \quad (37)
 \end{aligned}$$

Here  $k_2$  and  $k_3$  as illustrated in Fig. 11 are some unknown terms related to the cable and connector leakage, connection repeatability, drift, linearity and frequency error. And  $\Gamma_{e3}$  and  $\Gamma_{e2}$  are the so called equivalent source match terms of port 3 and port 2 respectively.

When the signal flow graph is used in power sensor calibration, recall that the power measured by a sensor  $P_m$  is actually  $|b_i|^2$  times calibration factor  $K$ , i.e.

$$P_m = |b_i|^2 \times K \quad (38)$$

where  $|b_i|^2$  is the power delivered to port  $i$  and  $K$  is the calibration factor.

Recall that in the measurement setup shown in Fig. 4, port 3 of the splitter is connected to a power sensor and meter, serving as a monitoring arm. Port 2 is a test arm, alternatively connected with standard and DUT power sensor. Power is measured at both port 2 and port 3.

When DUT is connected to the port 2, from equation (38) we can obtain the following ratio:

$$\frac{P_{DUT}}{P_{3DUT}} = \frac{|b_{2DUT}|^2 \times K_{DUT}}{|b_{3DUT}|^2 \times K_{3DUT}}$$

$P_{DUT}$  and  $P_{3DUT}$  are the powers measured at port 2 when DUT power sensor connected and that at port 3 respectively.  $K_{DUT}$  and  $K_{3DUT}$  are the calibration factors corresponding to the respective DUT sensor and sensor connected at port 3. Then

$$K_{DUT} = K_{3DUT} \times \frac{P_{DUT}}{P_{3DUT}} \times \frac{|b_{3DUT}|^2}{|b_{2DUT}|^2} \quad (39)$$

Alternatively connection with port 2 is replaced with the standard power sensor. Port 3 remains as the monitoring arm. With the changes made at port 2, the similar ratio can be

obtained,  $\frac{P_{Std}}{P_{3Std}} = \frac{|b_{2Std}|^2 \times K_{Std}}{|b_{3Std}|^2 \times K_{3Std}}$  and

$$K_{Std} = K_{3Std} \times \frac{P_{Std}}{P_{3Std}} \times \frac{|b_{3Std}|^2}{|b_{2Std}|^2} \quad (40)$$

Since  $K_{3Std} = K_{3DUT}$  for port 3 remains as the monitoring arm no matter port 2 connected with DUT or standard sensor, ratio of equation (39) and equation (40) yields

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{3DUT}} \times \frac{|b_{3DUT}|^2}{|b_{2DUT}|^2} \times \frac{P_{3Std}}{P_{Std}} \times \frac{|b_{2Std}|^2}{|b_{3Std}|^2} \quad (41)$$

Link to the transfer function derived in equation (37), and let  $a_x = b_2$  and  $a_y = b_3$

$$\frac{b_2}{b_3} = \frac{a_x}{a_y} = \left(\frac{k_2}{k_3}\right) \left(\frac{S_{21}}{S_{31}}\right) \left(\frac{1-\Gamma_r\Gamma_{e3}}{1-\Gamma_d\Gamma_{e2}}\right) \quad (42)$$

Further modifications are made to equation (41) to specify the transfer function for the alternate connections of the standard power sensor and the DUT power sensor to port 2 of the splitter while port 3 remains as the monitoring arm. Then:

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{3DUT}} \times \frac{P_{3Std}}{P_{Std}} \times \left|\frac{k_{2Std}}{k_{2DUT}}\right|^2 \times \left|\frac{1-\Gamma_{DUT}\Gamma_{e2}}{1-\Gamma_{Std}\Gamma_{e2}}\right|^2 \quad (43)$$

Compare the derived equation (43) from signal flow graph Fig. 11(b) with equation (29) derived from Fig. 4, the only difference is  $\left|\frac{k_{2Std}}{k_{2DUT}}\right|^2$  term. This term accounts for connector leakage, repeatability and so on. It is the Type A uncertainty shown in equation (5). The term is not explicitly expressed in the calibration equation (29). From different analytical method, we derive the same power sensor calibration equations.

### 5.4 Case study 2: Simultaneous comparison method analyses

Another way of using a resistive power splitter is the simultaneous comparison method. The reference standard and DUT power sensors to be calibrated are connected to the two output arms of the power splitter simultaneously. It seems an easy way of connection, no alternative connection of reference standard and DUT at port 2 of splitter as the above analyses. And there is application in some calibration laboratories.

With signal flow graph the transfer function can be easily obtained and reveal some findings. From equation (38), when DUT is connected to the port 2 and reference standard is connected to the port 3, we can obtain the following ratio:

$$\frac{P_{DUT}}{P_{3Std}} = \frac{|b_{2DUT}|^2 \times K_{DUT}}{|b_{3Std}|^2 \times K_{3Std}}$$

$P_{DUT}$  and  $P_{3Std}$  are the powers measured at DUT power sensor connected to port 2 and the standard power sensor connected to port 3 respectively.  $K_{DUT}$  and  $K_{3Std}$  are the calibration factors corresponding to the respective sensors. Then

$$K_{DUT} = K_{3Std} \times \frac{P_{DUT}}{P_{3Std}} \times \frac{|b_{3Std}|^2}{|b_{2DUT}|^2} \quad (44)$$

Substitute equation (42) and simplify,

$$K_{DUT} = K_{Std} \times \frac{P_{DUT}}{P_{3DUT}} \times \left| \frac{k_{3Std}}{k_{2DUT}} \right|^2 \times \left| \frac{S_{31}}{S_{21}} \right|^2 \times \left| \frac{1 - \Gamma_{DUT} \Gamma_{e2}}{1 - \Gamma_{3Std} \Gamma_{e3}} \right|^2 \quad (45)$$

Compare equation (45) with equation (43), correction for mismatch when using the simultaneous comparison method is more complicated than for the direct comparison transfer method in which one port is used to monitor the signal level and the power sensors to be compared are connected in turn to the other output port. It can be seen from equation (45) that the mismatches at both output ports of the power splitter are needed to be taken into account. Unless perfect symmetry can be assumed, it is necessary to determine for each output port the value of  $\Gamma_{e3}$  and  $\Gamma_{e2}$ . For this reason the simultaneous comparison method is not often used for measurement where the highest precision is required. However the method is quite convenient if mismatch errors are not to be corrected for, provided that errors due to asymmetry are eliminated by interchanging the two power sensors to be compared.

### 5.5 Case study 3: Extend the model to waveguide application

As proved by equation (43), equation (29) actually can be extended to use for any three port structure, not only the coaxial splitter based calibration system. For waveguide power sensor calibration, the directional coupler replaces splitter as the media components. The same equation then is used for WR15 (50 GHz - 75 GHz) sensor calibration. The system setup is illustrated in Fig. 12.

## 6. Source mismatch determination and uncertainty

As mentioned in the beginning of section 4, the splitter provides small value of effective source reflection coefficient  $\Gamma_{EG}$  at its output port 2 when used in a leveling or ratio measurement mode. And  $\Gamma_{EG}$  can be calculated from its complex S-parameters by equation (27):



$$\Gamma_{EG} = S_{22} - S_{21} \frac{S_{32}}{S_{31}}$$

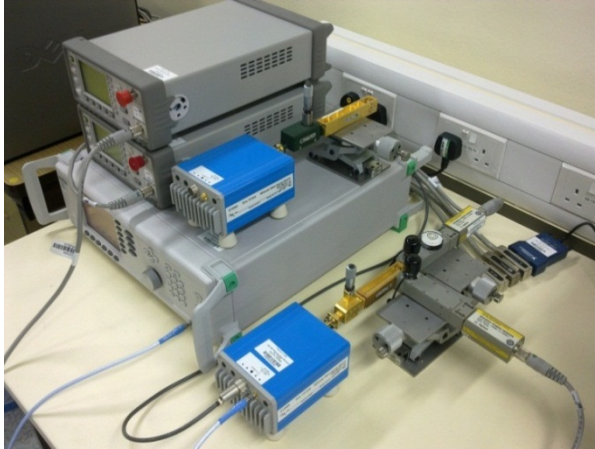


Fig. 12. Rectangular waveguide power sensor calibration system setup

One possible way of determining  $\Gamma_{EG}$  is to measure the splitter's complex S-parameters and calculate its value from equation (27). The measurement uncertainty in the S-parameters is then evaluated (EURAMET, 2011). Reference (Ridler and Salter, 2001) presented the law of propagation of uncertainty using matrix notation, treating the complex quantities with real and imaginary evaluation.

Since the calculation method from S-parameters is sensitive to small measurement errors, several different measurement methods were proposed, such as the "passive open circuit" and "active open circuit" method (Moyer, 1987), and the direct calibration method (Juroshek, 1997). In the direct calibration method, the splitter is connected through ports 1 and 3 to a VNA. This effectively gives a new one port VNA at splitter port 2. This new VNA is calibrated using a one port calibration algorithm, e.g. short-open-load. The  $\Gamma_{EG}$  is then obtained as one of the three one-port VNA error terms. References (Rodriguez, 2000; Yhland & Stenarson, 2007) assessed the measurement uncertainty and traceability in power splitter effective source reflection coefficient. Reference (Furrer, 2007) compared direct calibration method with the calculation method. It seems that the similar results were obtained.

## 7. Conclusion

From simple direct comparison transfer method, to coaxial splitter based direct comparison transfer, and then to the general models with signal flow graph analyses, the chapter has meticulously discussed RF and microwave power sensor calibration methods. The models and equations provided target for practical usage. The examples and case studies have shown the practical applications. The traceability and measurement uncertainty with GUM and MCM have provided in details. The calibration models and methods described are useful for the coaxial and waveguide power sensor calibrations. The general models can be further developed for different case analyses in future studies.

## 8. Acknowledgement

The authors would like to thank the software and operation of their colleagues and students in the relative projects involved in the chapter.

## 9. References

- Agilent, (2003). Application Note AN1449-1/2/3/4, *Fundamentals of RF and Microwave Power Measurements*, 2003-2011, Retrieved from: <[www.agilent.com](http://www.agilent.com)>
- Brunetti, L. and Vremera, E. (2003). A new microcalorimeter for measurements in 3.5-mm coaxial line, *IEEE Transactions on Instrumentation and Measurement*, vol. 52, No. 2, pp. 320-323, Apr. 2003
- Bryant, G. H. (1993). *Principles of Microwave Measurements*, revised Edition, pp.20-38, published by Peter Peregrinus Ltd., ISBN 0 86341 296 3, London, UK
- Clague, F. R. (1995). *A Calibration Service for Coaxial Reference Standards for Microwave Power*, NIST Technical Note 1374, May 1995
- Crowley, T. (2006). Microwave Power Measurement, NIST/ARFTG Measurement short course, 2006
- Cui, X. and Crowley, T. P. (2011). Comparison of experimental techniques for evaluating the correction factor of a rectangular waveguide microcalorimeter", *IEEE Transactions on Instrumentation and Measurement*, Vol.60, No.7, pp.2690-2695, July 2011
- Engen, G. F. (1992). *Microwave circuit theory and foundations of microwave metrology*, published by Peter Peregrinus Ltd, ISBN 0 86341 287 4, London, UK
- EURAMET/cg-12/v.02, (2011). *Guidelines on the Evaluation of Vector Network Analyzers (VNA)*, Retrieved from: <<http://www.euramet.org/index.php?id=calibration-guides>>
- Famton, A. (1990). *Radio Frequency & Microwave Power Measurement*, published by Peter Peregrinus Ltd, ISBN 0 86341 120 7, London, UK
- Furrer, J. (2007). Comparison of 50 GHz Splitter Source Match Measurement Methods (Juroshek / Russell-hp), RF-Power Lab, METAS, 1.11.2007
- Ginley, R. (2006). A direct comparison system for measuring radio frequency power (100 kHz to 18 GHz), *MEASURE*, pp.46-49, Vol.1, No.4, December 2006
- Joint Committee for Guides in Metrology (JCGM) 100:2008. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement*. Copyright of the document is shared jointly by the JCGM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML). Retrieved from: <<http://www.iso.org/sites/JCGM/GUM-introduction.htm>>
- JCGM 101:2008. *Evaluation of measurement data – Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement – Propagation of Distributions using a Monte Carlo method'*. Copyright of the document is shared jointly by the JCGM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML). Retrieved from: <<http://www.iso.org/sites/JCGM/GUM-introduction.htm>>
- JCGM 200:2008. *International Vocabulary Of Metrology – Basic and General Concepts and Associated Terms (VIM)*, Copyright of the document is shared jointly by the JCGM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML). Retrieved from <<http://www.iso.org/sites/JCGM/VIM-introduction.htm>>
- Johnson, R. A. (1975). Understanding Microwave Power Splitters, *Microwave Journal*, pp.49-56, December 1975
- Juroshek, J. R. (1997). A direct calibration method for measuring equivalent source mismatch, *Microwave Journal*, pp. 106-118, Oct. 1997

- Juroshek, J. R. (2000). NIST 0.05-50 GHz direct comparison power calibration system, *Conference on Precision Electromagnetic Measurements (CPEM2000)*, pp.166-167, 14-19 May 2000, Sydney, Australia
- Mial, J. (2007). Chapter 15: RF power measurement, In: *Microwave Measurements*, 3rd edition, Edited by R.J. Collier and A.D. Skinner, pp.330-348, published by The Institution of Engineering and Technology, ISBN 978-0-86341-735-1, London, UK
- Moyer, R. D. (1987). Techniques for Measuring the effective source reflection coefficient of two-resistor power splitters, *IEEE Transactions on Instrumentation and Measurement*, pp.23-28, Vol.IM-36, No.1, March 1987
- Oldfield, L.C. (1989). Chapter 7, Power measurement, In: *Microwave Measurement*, Edited by Bailey, A.E., second edition, pp. 105-131, published by Peter Peregrinus Ltd, ISBN 0 86341 184 3, London, UK
- Ridler, N. M. and Salter, M. J. (2001). Propagating S-parameter uncertainties to other measurement quantities, *58<sup>th</sup> Automatic RF Techniques Group (ARFTG) conference*, November 2001
- Ridler, N. M. and Salter, M. J. (2002). An approach to the treatment of uncertainty in complex S-parameter measurements, *Metrologia*, pp.295-302, Vol.39, No.3, 2002
- Rodriguez, M. (2000). Towards Traceability of the direct calibration method to the cal kit standards, *Conference on Precision Electromagnetic Measurements (CPEM2000)*, pp.636-637, 14-19 May 2000, Sydney, Australia
- Rumfelt, A. Y. and Elwell, L.B. (1967). Radio Frequency Power Measurements, *Proceedings of the IEEE*, pp.837-850, Vol.55, No.6, June 1967
- Shan, Y., Chua, S.W., Neo H. and Wu, T. (2008). A Direct Comparison Transfer Microwave Power Sensor Calibration System, *Conference on Precision Electromagnetic Measurements (CPEM08)*, pp 512-513, 8-13 June 2008, Colorado USA
- Shan, Y., Chua, S.W., Brunetti, L., Oberto, L. and Sellone, M. (2010a). Bilateral Comparison Between NMC And INRIM On Microwave Power Sensor Using Type N And 3.5 mm Connectors, *Conference On Precision Electromagnetic Measurements (CPEM2010)*, pp.736-737, 13-18 June, 2010, Daejeon, Korea
- Shan, Y., Chua, S.W. and Yan, Y.K. (2010b). Development of A 50 GHz Coaxial Direct Comparison Transfer Microwave Power Sensor Calibration System At NMC, *Conference On Precision Electromagnetic Measurements (CPEM2010)*, pp.738-739, 13-18 June, 2010, Daejeon, Korea
- Shimaoka, K., Shida, M. and Komiyama, K. (2006). Source reflection coefficient measurements of the power reference of power meters, *CPEM2006*, pp.664-665
- Torok, A., Janik, D., Peinelt, W., Stumpe, D., and Stumper, U. (2001). Efficient Broadband method for equivalent source reflection coefficient measurements, *IEEE Transactions on instrumentation and measurement*, pp.361-363, Vol.50, No. 2, April 2001
- Weidman, M.P. (1996). *Direct Comparison Transfer of Microwave Power Sensor Calibration*, NIST Technical Note 1379, January 1996
- Wong, K. (2002). Power Sensor Calibration and Uncertainties, *60<sup>th</sup> ARFTG Conference*, December 2002
- Yhland, K. and Stenarson, J. (2007). Measurement Uncertainty in Power Splitter Effective Source Match, *IEEE Transactions on Instrumentation and Measurement*, pp.669-672, Vol.56, No.2, April 2007

# Fast Method for Frequency Measurement by Rational Approximations with Application in Mechatronics

Daniel Hernandez-Balbuena<sup>1</sup>, Oleg Sergiyenko<sup>2</sup>,  
Patricia L. A. Rosas-Méndez<sup>1</sup>, Vera Tyrsa<sup>3</sup> and Moises Rivas-Lopez<sup>2</sup>

<sup>1</sup>*Engineering Faculty, Autonomous University of Baja California*

<sup>2</sup>*Engineering Institute of Autonomous University of Baja California*

<sup>3</sup>*Polytechnic University of Baja California, Mexicali  
México*

## 1. Introduction

In Automatic Control, Mechatronics and Robotics technical tasks is desirable to have information about certain parameters to control, supervision or fault detection in a system. This information can be estimates by direct measurement of particular variables or using special devices capable of observing the state of system under control. If the first option is selected a physical sensor is required.

One commonly selected option are the Frequency Domain Sensors (FDS), this devices converts the desired parameter into a square wave with a frequency or period proportional to physical quantity under measurement. In order to reduce time to obtain information for control a system with sufficient quality and good reliability, a high performance method for frequency measurement is desirable.

Historically, many analog and digital frequency measurement techniques have been proposed. In a basic digital measurement technique, the zero crossings of a signal are detected and a square wave is formed, representing transitions between the binary logic levels low and high. Selected digital logic state is detected and counted, and a measure of frequency is determined by the number of complete cycles occurring in the square waveform during a fixed time interval, determined by the counter's time base [1]. This method can be classified as the classical method and their main error source is the  $\pm 1$  error count derived from the relative timing of gate and signal, which means that the resolution is 1 Hz during a 1s gate time for all input signal frequencies [2]. For allow high resolution frequency measurements gate time larger than 1s should be selected.

In reciprocal counting measurement techniques, the gate time is determined by electronic detection of two kinds of same phase difference situations between two pulsed signals with

different frequencies [3], or by electronic detection of two coincident pulses of two regular independent pulse trains [5 -6]. In these methods the quantization error ( $\pm 1$  count error) can be overcome satisfactorily [2-3,6]. But, in [3] a high distinguishability analog circuit for phase coincidence detection is required and in [5] the relative methodical error is pulse width dependent for random selection of stop measurement pulse coincidence and, is experimentally probed that relative methodical error can be reduced by two or three orders of magnitude than frequency meters based on classical method [6].

Continuous time stamping principle change the scenario in frequency measurement, because in each measurement has not a defined start (= start trigger event), and a stop (= stop trigger event) plus a dead-time between measurements to read out and clear registers, do interpolation measurements and prepare for next measurement [2]. In this technique Linear regression using the least-squares line fitting is used because for a one-second frequency measurement in a fast processing counter could contain hundreds or thousands of paced time-stamped events, not just a start event plus a stop event [2]. But fast digital circuits are needed to implement this technique.

However, a fast method for frequency measurement base pulse coincidence principle and rational approximations was proposed and, was shown that under a novel numerical condition for detect the stop trigger event measurement resolution is improved. Instrumental errors are caused only by the reproducibility of the reference frequency and relative measurement error is comparable to the reproducibility of reference oscillator. [7]. Simple digital circuits are needed to practical implementation of this technique.

## 2. Frequency measurement based on the pulse coincidence principle

In the past, pulse coincidence principle has been used for frequency measurement of electrical signals [4-6]. In this measurement method, a desired frequency is measured by comparing it with a standard frequency. The zero crossings of both frequencies are detected, and a narrow pulse is generated at each crossing. Then, two regular independent pulse trains are generated. The desired and standard trains of narrow pulses are compared for coincidence. This is made with an AND-gate, and then a coincidence pulse train is generated. Coincident pulses can be used as triggers to start and stop a pair of digital counters (start and stop events). See Figure 1.

The standard and desired pulse trains are applied to other counters and a measure of the desired frequency is obtained by multiplying the known standard frequency by the ratio between the desired count and the standard count obtained in the two digital counters [5]. And basically, measurement error reduction depends on the ability to detect a pair of *best* coincidences consecutively between a large number of these, for a given measurement time. In other words, measurement error reduction depends on reduction of comparison error of time intervals  $n_0 T_0$  and  $n_X T_X$  by selection of adequate coincidence pulses [6].

Consider  $f_x$  as the desired or unknown frequency and  $f_o$  as the standard frequency, and  $T_x = 1/f_x$  and  $T_o = 1/f_o$  are they periods respectively. In Figure 1,  $S_X(t)$  and  $S_o(t)$  are the unknown and standard trains of narrow pulses,  $\tau$  is the pulse width on both trains and,  $S_X(t) \& S_o(t)$  is the irregular pulse coincidence train. Then,  $n_0$  and  $n_X$  are integer numbers and represent the amount of whole periods between selected measurement start and stop events (see figure 1).

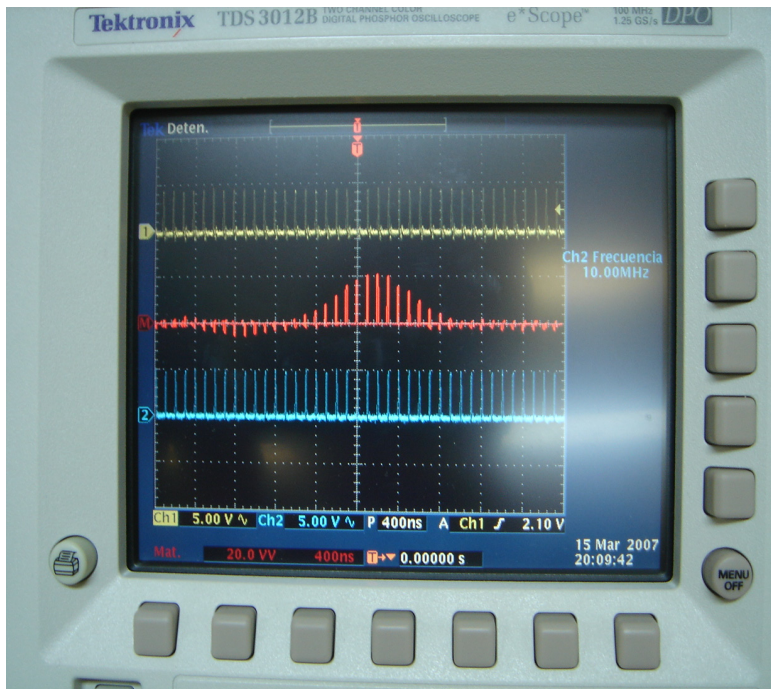
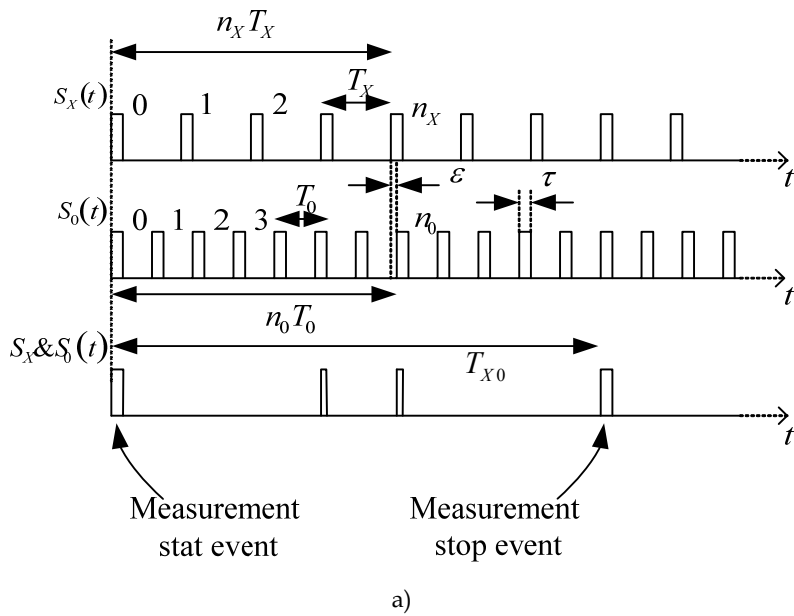


Fig. 1. a) Pulse coincidence principle for frequency measurement, b) Practical view of coincidences process (oscilloscope screenshot).

In [6] by possible selection of *partial* coincidences, error of comparison of time intervals  $n_0T_0$  and  $n_XT_X$  has been reduced to the duration of the coinciding pulses and mathematically expressed by

$$|n_XT_X - n_0T_0| \leq 2\tau. \quad (1)$$

Form (1), relative measurement error for a single measurement can be expressed as follows

$$\beta_X = \frac{\left| f_X - \frac{n_X}{n_0} f_0 \right|}{f_X} \leq \frac{2\tau}{n_0T_0} \approx \frac{2\tau}{t_m}, \quad (2)$$

where  $t_m$  is the measurement time.

It is experimentally known, that for a standard frequency  $f_0 = 1 \times 10^6$  Hz from a thermostatic quartz generator with relative short-term instability does not exceed  $10^{-8}$  and using a pulses width  $\tau \approx 7 \times 10^{-9}$  s, root-mean-square (RMS) error  $s_X = 14.3 \times 10^{-6}$  Hz for  $f_X = 1 \times 10^3$  Hz at 50 observations during total measurement time  $t_m \leq 1$  s. And  $s_X = 1.79 \times 10^{-3}$  Hz for  $f_X = 1 \times 10^6$  Hz under same measurement conditions [6].

## 2.1 functioning and uncertainty limitations

Let us consider two trains of narrow pulses with period  $T_X$  and  $T_0$  with and pulse width  $\tau$  respectively, generated by detection of zero crossings of two sinusoidal signal of frequencies  $f_0$  and  $f_X$ . Suppose that  $T_0$  is a known parameter and  $T_X$  is unknown and, both pulse trains start in phase, i.e. a time shift is 0.

For an appropriate selection of the pulse width of two regular independent pulse trains, periodic perfect coincidences of these pulses are observed in time axis [13]. Repetition period of perfect coincidences is  $T_{X0}$ , in Fig. 1a). A Practical view of coincidences process is presented in Fig 1 b).

For frequency measurement, the time intervals  $n_0T_0$  and  $n_XT_X$  are compared, where  $n_0$  is the amount of periods  $T_0$  in the measurement time and  $n_X$  is the amount of periods  $T_X$  in the same time interval.

Measurement time can be defined by the time interval between the first one pulse of coincidence (start event) after beginning the measurement process, and by any other following pulse of coincidence (stop event). As it were mentioned in the previous section,  $n_0$  and  $n_X$  are the counts of pulses obtained in two digital counters.

According to [12], pulse coincidence occurs when

$$|n_XT_X - n_0T_0| \leq \varepsilon \quad (3)$$

where  $\varepsilon$  is the acceptable tolerance (reasonable error value between time intervals  $n_0T_0$  and  $n_XT_X$  less to pulse with and dependent on the quality of electronic circuits used).

To find values of  $n_0$  and  $n_X$  which define the appropriated coincidence it is useful to expand  $T_X/T_0$  as simple continued fractions. This is evident rewriting (3)

$$\left| \frac{T_X}{T_0} - \frac{n_0}{n_X} \right| \leq \frac{\varepsilon}{n_X T_0}. \quad (4)$$

Left side of equation (3) represents the approximation of  $T_X/T_0$  using rational numbers and right side is an approximation condition.

For frequency measurement, in a view of  $f_0 = 1/T_0$ ,  $f_X = 1/T_X$ , we can write

$$\left| f_X - \frac{n_0}{n_X} f_0 \right| \leq \frac{\varepsilon f_X f_0}{n_0}. \quad (5)$$

In (5),  $f_X$  is the hypothetical true value of the unknown frequency and  $f_0 n_0/n_X$  is the frequency value obtained by the measurement. Then, dividing both parts of (5) in  $f_X$  and taking to account  $f_0 = 1/T_0$ , relative error of measurement (frequency offset)  $\beta$  can be expressed by

$$\beta = \left| f_X - \frac{n_0}{n_X} f_0 \right| / f_X \leq \frac{\varepsilon}{n_0 T_0}. \quad (6)$$

We can see in equation (6) that relative error of measurement is limited by the ratio between the acceptable tolerance of the error of comparison between the time intervals  $n_0 T_0$  and  $n_X T_X$  and, the time interval  $n_0 T_0$ . Value of  $n_0 T_0$  is approximately the measurement time.

## 2.2 Numerical stop condition of measurement

In frequency measurement,  $n_0$  and  $n_X$  are independent counter counts obtained in two digital counters, so they are properly integer numbers. An integer numbers ratio, like the involved in the frequency value obtained by the measurement, is possible to investigate under number theory laws. Let note and briefly explain some of them, especially Euclidean algorithm.

### 2.2.1 Number theoretic preliminaries

Let us to assume without loss generality that  $T_X > T_0$ , from the division algorithm we can write

$$T_X = a_0 T_0 + \Delta t_0 \quad T_0 > \Delta t_0 \geq 0 \quad (7)$$

$$T_0 = a_1 \Delta t_0 + \Delta t_1 \quad \Delta t_0 > \Delta t_1 \geq 0 \quad (8)$$

$$\Delta t_0 = a_2 \Delta t_1 + \Delta t_2 \quad \Delta t_1 > \Delta t_2 \geq 0 \quad (9)$$

$$\vdots \qquad \qquad \qquad \vdots$$



$$\Delta t_{i-2} = a_i \Delta t_{i-1} + \Delta t_i \quad \Delta t_{i-1} > \Delta t_i \geq 0 \tag{10}$$

⋮

$$\Delta t_{n-2} = a_n \Delta t_{n-1} + \Delta t_n \quad \Delta t_{n-1} > \Delta t_n \geq 0 \tag{11}$$

where the  $a_i$  is the  $i$ th partial quotients for each case and  $\Delta t_i$  is the  $i$ th remainder, with  $i = 1, 2, 3, \dots, n$ . With  $a_i \geq 1$ ,  $\Delta t_i$  is a decreasing sequence for  $i \geq 0$ .

Each remained obtained in the division step of Euclidean algorithm be could be interpreted as a distance [11], defined by

$$|Q_i T_X - P_i T_0|_0 = \Delta t_i. \tag{12}$$

where  $P_i$  and  $Q_i$  are the numerator and denominator of the  $i$ th convergent of the continued fractions to  $T_X/T_0$  defined recursively as [12]

$$P_i = a_i P_{i-1} + P_{i-2} \tag{13}$$

$$Q_i = a_i Q_{i-1} + Q_{i-2} \tag{14}$$

for arbitrary  $i \geq 2$ , and

$$\begin{aligned} P_0 &= a_0, & Q_0 &= 1, \\ P_1 &= a_0 a_1 + 1, & Q_1 &= a_1. \end{aligned}$$

Then, from (12) each remainder  $\Delta t_i$  is the absolute difference between the time intervals  $Q_i T_X$  and  $P_i T_0$ .

On the other hand,  $T_0$  can be expressed in terms of two consecutive remainders [11] using the following expression:

$$T_0 = Q_i \Delta t_{i-1} + Q_{i-1} \Delta t_i. \tag{15}$$

A similar expression can be derived for  $T_X$

$$T_X = P_i \Delta t_{i-1} + P_{i-1} \Delta t_i. \tag{16}$$

Supposing that  $n$  is the number of steps in the Euclidean algorithm to obtain greatest common divisor of  $T_X$  and  $T_0$ . Then last remained  $\Delta t_n = 0$  and time interval  $\Delta t_{n-1}$  is the greatest common divisor of both periods  $T_X$  and  $T_0$ , in the consecutive division expressed in equations (7) to (11). Because the greatest common divisor is: the last nonzero remainder in this sequence of divisions.

Assuming that  $\Delta t_{n-1}$  is greatest common divisor of both periods  $T_X$  and  $T_0$ , we can write

$$T_0 = Q_n \Delta t_{n-1} \tag{17}$$

$$T_X = P_n \Delta t_{n-1}. \tag{18}$$

Expressing (10) in terms of (15) and (16) it is evident in (17) than step  $n$  is total equality point for both time intervals

$$|Q_n P_n \Delta t_{n-1} - P_n Q_n \Delta t_{n-1}| = 0. \quad (19)$$

In frequency measurement, this term expressed in several forms have a mathematical mean of least common multiple, and practical mean of time interval  $T_{X0}$  (see Fig. 1) expressed by:

$$T_{X0} = \frac{T_X T_0}{\Delta t_{n-1}} = P_n Q_n \Delta t_{n-1} \quad (20)$$

is the condition for periodic perfect coincidences of pulses (see Fig. 1). Assuming to  $T_{X0}$  the measurement time in frequency measurement, from (1) and Fig. 1

$$|n_X T_X - n_0 T_0| = 0. \quad (21)$$

and each time intervals  $n_0 T_0$  and  $n_X T_X$  are equal to  $T_{X0}$ . Then,

$$n_0 T_0 = P_n Q_n \Delta t_{n-1}, \quad (22)$$

$$n_X T_X = P_n Q_n \Delta t_{n-1}. \quad (23)$$

Now, product of two numbers  $ab = c$  can be considered as the sum  $a + a + a + \dots + a$  in which the number of summands is equal to  $b$  or as the sum  $b + b + b + \dots + b$  in which the number of summands is equal to  $a$ .

Then, equations (22) and (23) can be rewritten using (17) and (18):

$$n_0 Q_n \Delta t_{n-1} = P_n Q_n \Delta t_{n-1},$$

$$n_X P_n \Delta t_{n-1} = P_n Q_n \Delta t_{n-1}.$$

Expressions (22) and (23) have a reason only when

$$n_0 = P_n \quad (24)$$

and

$$n_X = Q_n. \quad (25)$$

### 2.2.2 Stop condition of measurement

For small measurement time (less than or equal a 1 s) is evident from equation (4) that: order of magnitude of  $\Delta t_{n-1}$  must be of same order of magnitude that the expected relative error of measurement  $\beta$ . Then, according to before mentioned, we propose that an acceptable tolerance in (1) is  $\varepsilon = \Delta t_{n-1}$ .

Assuming decimal notation for both periods  $T_X$  and  $T_0$ , under the conditions  $T_X < 1$  and  $T_0 < 1$ , and assuming reference period can be expressed as  $T_0 = 1 \times 10^{-s}$ , then the greatest common divisor  $\Delta t_{n-1}$  is

$$\Delta t_{n-1} = (T_X, T_0) = \frac{1}{10^r} (A, 10^{r-s}) \quad (26)$$

where  $A$ ,  $r$ ,  $s$  are integer numbers with  $r > s$ ,  $r$  is the exponent associate to expected the order of magnitude of  $\beta$ ,  $r-s$  is the difference between the expected order of magnitude  $\beta$  and the order of magnitude of the time period of the standard.

On the other hand, according with equation (25) the number of time intervals  $T_X$  necessaries to stop the measurement process is  $Q_n$ , and form (15)

$$Q_n = \frac{T_0}{\Delta t_{n-1}}. \quad (27)$$

If  $A$  and  $10^{r-s}$  in equation (26) are mutually prime then  $\Delta t_{n-1} = 10^{r-s}$  and

$$Q_n = 10^{r-s} \quad (28)$$

and, if they are not, then  $\Delta t_{n-1} = a/10^r$  whit  $a$  integer number and

$$Q_n = 10^{r-s} / a. \quad (29)$$

In both cases  $1/10^r$  is a common divisor of both periods  $T_X$  and  $T_0$ .

Then, from equations (16), (24) and (25), the condition that satisfies (19) is

$$n_X = 10^{r-s}. \quad (30)$$

This is the numeric condition that we propose to stop the measurement process and is easy to implement with basics digital circuits.

A novel fast method to frequency measurement with application in mechatronics and telecommunication is based on this numerical stop condition is presented in [7-9]. Resolution improvement in frequency domain sensors is allowed in automotive applications [10] with this method and is applied in precise optical scanning and structural health monitoring [19-21]

### 3. Simulation

In the simulation two pulse trains of unitary amplitude are generated using a computational algorithm sampling independent [7]. The value of reference frequency was accepted as  $f_0 = 1 \times 10^7$  Hz. The hypothetical value of unknown frequency is  $f_X = 5878815.277629991$  Hz, and is a result of the accepted value of the period  $T_X = 1.701023 \times 10^{-7}$  s. The value of pulse width in both pulse trains is accepted as  $\tau = 1.5 \times 10^{-9}$  s.

In this case, is evident that  $T_X$  and  $T_0$  are mutually prime numbers and have the common denominator  $\Delta t_{n-1} = 1 \times 10^{-13}$ .

Simulation algorithm provided continuous formation of the segments  $n_0 T_0$  and  $n_X T_X$  and compares the magnitude of their difference with parameter  $2\tau$ . When the value of the

specified difference was less than  $2\tau$  on corresponding steps of simulation, it was identified like a coincidence of pulses and the integer numbers  $n_0$  and  $n_X$  are stored.

The unknown frequency is calculated using  $f_{Xm} = n_X f_0 / n_0$  and frequency relative error is obtained using  $(f_X - f_{Xm}) / f_X$ , both results are stored also.

Simulation results are partially presented in the Table 1 and, frequency relative error calculated (non absolute value) is presented in Fig. 2 for a simulation time of 0.2 s.

The simulation process star in  $n_0 = 0$  and  $n_X = 0$ , and the best approximation is selected (in this case) using the condition,  $n_X = 1 \times 10^6$ .

Table 1 represents an interesting fact. For thousands of data we have the same uncertainty range  $10^{-13}$ , as for first and third rows. And only when  $n_X$  takes a form of 1 with six zeros (in this case, second row of Table 1) we are getting up to  $10^{-17}$ .

In Fig. 2, we can see a global convergence to zero of frequency relative error. An alternated convergence and a non monotone decreasing characteristic are evident. However, we can identify in the graphic a point where  $\beta$  is minimum for an approximated time of 0.17 s.

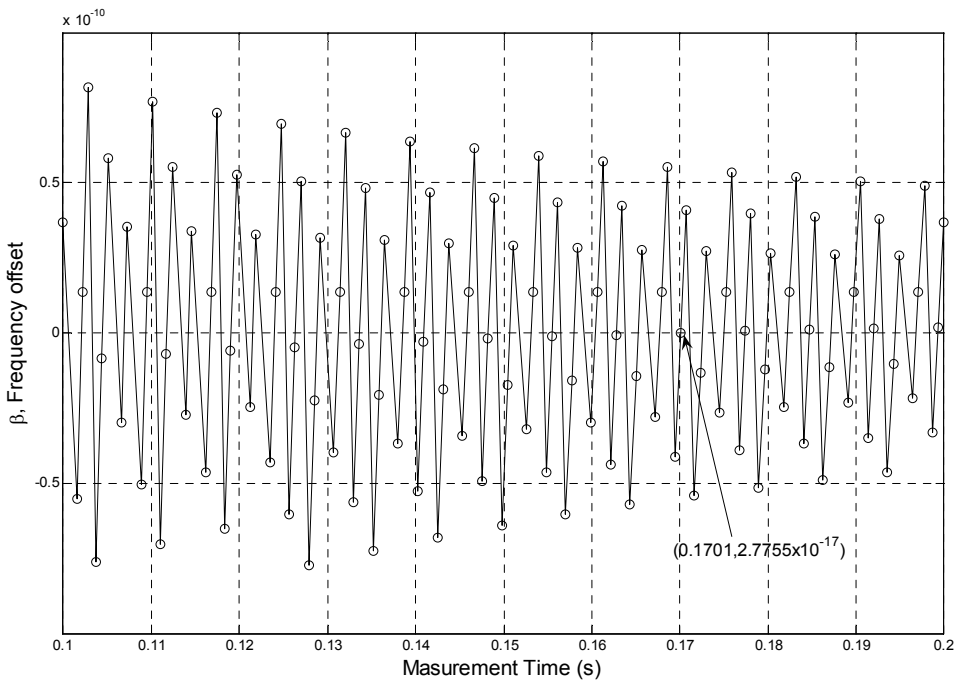


Fig. 2. Frequency offset from the simulation process

Frequency offset from computational selection of best coincidences, obtained under condition (1) with  $\varepsilon = 1 \times 10^{-12}$  is presented in Fig. 3. In this graphic, we can observe a convergence by the left and a divergence by the right around 0.17 s (first point in the Fig. 3).

This condition is repeated with time, and we can see five points where absolute value of  $\beta$  is minimum, for a simulation time of 1 s. In each mentioned points the condition expressed in (30) is fulfilled.

$n_X$	$n_0$	$ n_X T_X - n_0 T_0 , s$	$f_{Xm}, Hz$
957087	1628027	$1.00 \times 10^{-13}$	5878815.277633602
1000000	1701023	$2.78 \times 10^{-17}$	5878815.277629991
1042913	1774019	$1.00 \times 10^{-13}$	5878815.277626677

Table 1. Simulation results of frequency measurement process.

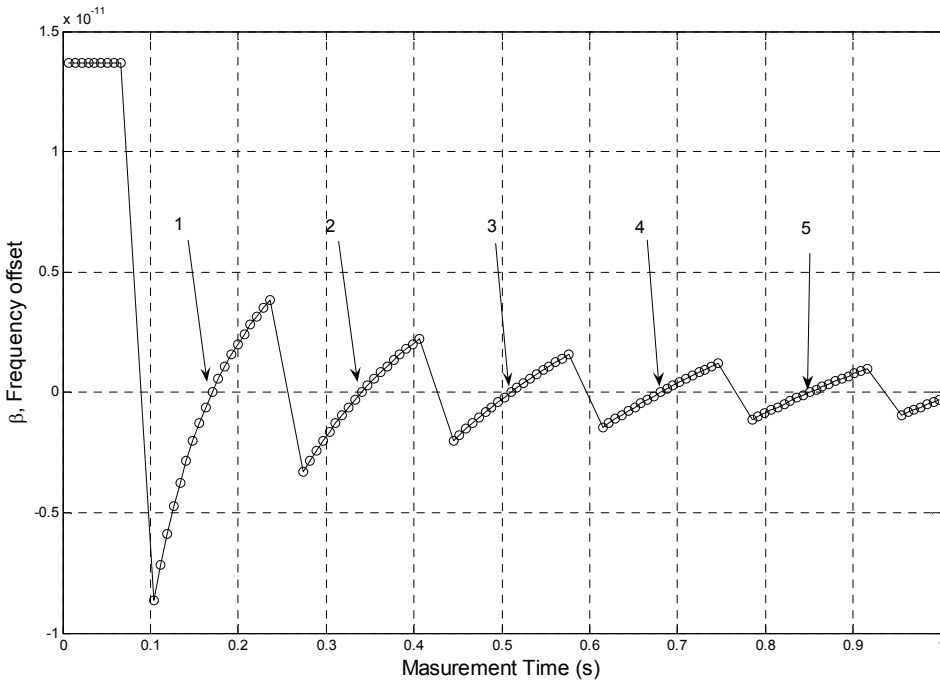


Fig. 3. Frequency offset from the simulation process for 1 s.

### 3.1 Jitter effect in frequency measurement

In order to evaluate the jitter effect on the non-electrically detectable stop event for frequency measurement method based on the direct comparison of two regular independent trains of narrow pulses and rational approximations. Deterministic and random components of jitter are modeled and, are added in both pulse trains one deterministic jitter component and random jitter in each case. Simulation results are presented when both pulse trains start in phase and when start with a phase shift.

Timing jitter (henceforth referred to as jitter) is defined as short-term non-cumulative variations of the significant instants of a signal from their ideal positions in time [18].

For modeling, Total Jitter (TJ) consists of two components: Deterministic Jitter (DJ) and Random Jitter (RJ) [14]. In time domain, TJ is the sum of the RJ and DJ components [15]. RJ is characterized by a Gaussian distribution. It has been shown that it is theoretically unbounded in amplitude.

DJ consists of several components caused by different and mostly physically-based phenomena, such as electronic interference, cross-talk and bandwidth limitation. All DJ subcomponents have a bounded peak-to-peak value that does not increase when more measurement samples are taken [15].

Deterministic jitter has four components: duty cycle distortion (DCD), intersymbol interference (ISI), periodic jitter (PJ) and bounded uncorrelated jitter (BUJ).

DCD and ISI are referred as data correlated jitter, while PJ and BUJ are referred as data uncorrelated jitter. RJ is unbounded and uncorrelated [15].

### 3.1.1 Random Jitter (RJ)

Random Jitter RJ is caused the common influence of a large number of very small independent contributor or various device-originated noise sources (such as thermal and flicker noise). By the central limit theorem, the distribution of a large number of uncorrelated noise sources approaches a probability Gaussian distribution and is given by [14]

$$J_{RJ}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x^2}{2\sigma^2}\right)} \quad (31)$$

where  $\sigma$  is the standard deviation of the jitter distribution or the RMS value, and  $J_{RJ}$  is the probability that leading edge (or trailing edge) will occur at time  $x$ , where  $x$  is the deviation from the mean value of the time reference point (time point related to 50% amplitude point on pulse edge). In Fig. 4a), is shows the histogram for random jitter.

### 3.1.2 Periodic Jitter (PJ)

Periodic jitter denotes periodical timing deviation from the ideal position of a signal that repeat in time, is typically uncorrelated to the data rate or the clock frequency [14]. Electromagnetic interference and crosstalk from some clock line can cause periodic jitter.

The mathematic model of PJ consists of a sum of cosine functions with phase deviation, modulation frequency, and peak amplitude. The model is given by

$$PJ_T = \sum_{i=0}^n a_i \cos(\omega_i t + \theta_i) \quad (32)$$

where  $PJ_T$  denotes the total periodic jitter,  $n$  is the number of cosine components,  $a_i$  is the amplitude in units of time in each tone,  $\omega_i$  is the angular frequency of the corresponding modulation,  $t$  is the time, and  $\theta_i$  is the corresponding phase [15].

Sinusoidal jitter in time domain produces a probability distribution function given by (defining time zero as the center of the distribution)

$$J_{PJ_i}(x) = \begin{cases} \frac{1}{\pi\sqrt{a_i^2 - x^2}} & |x| \leq a_i \\ 0 & \text{otra} \end{cases} \quad (33)$$

where  $2a$  is the peak-to-peak width of periodic jitter [14]. In Fig. 4b) is shows the histogram for sinusoidal periodic jitter with added random jitter.

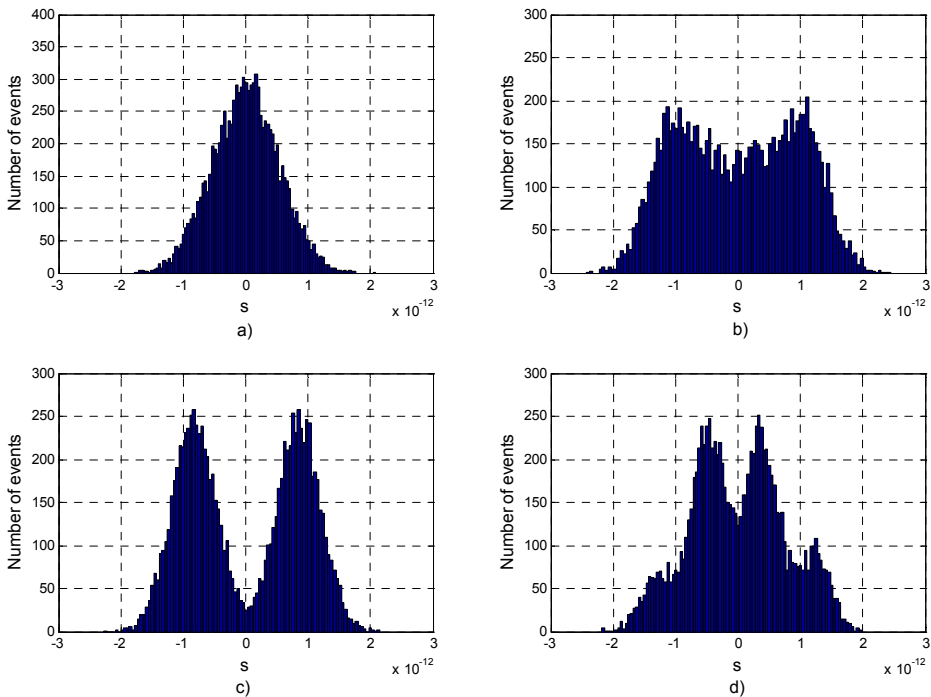


Fig. 4. Histogram for a) Random jitter (RJ), b) Sinusoidal periodic jitter with added random jitter (PJ+RJ), c) duty cycle distortion with added random jitter (DCD+RJ) y d) data dependent jitter with added random jitter (DDJ+RJ).

### 3.1.3 Duty Cycle Distortion (DCD)

Duty Cycle Distortion is often also called pulse width distortion [14], is deviation in the duty cycle value from the ideal value, this equates to a deviation in bit time between a 1 bit (logic 1) and a 0 bit (logic 0). DCD can have several sources. The most common are threshold level offsets and differences in the rising and falling edge characteristics [14].

DCD yields a binomial distribution consisting of two sharp peaks of equal height, unless one separates rising and falling transitions in the measurement. Theoretically those peaks are Dirac delta functions, but in practice random jitter and limited measurement resolution always produce peaks of finite height and finite width. The analytic equation for DCD distribution is the sum of delta functions [14]:

$$J_{DCD}(x) = \frac{\delta(x-a)}{2} + \frac{\delta(x+a)}{2} \quad (34)$$

where  $2a$  is the peak-to-peak width of the DCD. In Fig. 4c) is shows the histogram for duty cycle distortion with added random jitter.

### 3.1.4 Data Dependent Jitter (DDJ)

Data dependent jitter describes timing errors that depend on the preceding sequence of data bits [14]. DDJ is a predominant form of DJ caused by bandwidth limitations of the system or electromagnetic reflections of the signal [16-17]. Since there is always only a limited number of different possible patterns in a data stream of limited length, data dependent timing errors always produce a discrete timing jitter, theoretically DDJ distribution is the sum of two or more delta functions [14]:

$$J_{DDJ}(x) = \sum_{i=1}^N \{p_i \delta(x - t_i)\} \quad (35)$$

where  $\sum_{i=1}^N p_i = 1$   $N$  is number of distinct patterns,  $p_i$  is the probability of the particular pattern occurring, and  $t_i$  is the timing displacement of the edge following this pattern. In Fig. 4d is shows the histogram for data dependent jitter with added random jitter.

In the simulations, two pulse trains of unitary amplitude are generated. The value of reference frequency was accepted as  $f_0 = 1 \times 10^7$  Hz. The hypothetical value of unknown frequency  $f_X = 5878815.277629991$  Hz is a result of the accepted value of the period  $T_X = 1.701023 \times 10^{-7}$  s, and value of pulse width in both pulse trains is  $\tau = 1.0 \times 10^{-9}$  s. The RJ model used is Gaussian distributed with RMS value 0.7 ps. The PJ model is a single-tone sinusoidal with frequency 5 MHz and peak-to-peak value of 10 ps, same peak-to-peak value is assumed in DDJ and DCD jitter models. A component of random jitter is added to the last tree models to generate the jitter models PJ+RJ, DDJ+RJ and DCD+RJ. Such jitter values are selected because of the most typical values for such pulse trains according to [15]. Then the jitter components modeled are applied to the time reference points of each narrow pulse on the two pulse trains. In

## 3.2 Jitter simulation results

In the first computational experiment is assumed the same magnitude of the jitter in the oscillator under measurement and references oscillator on all models. The experimental results when the two pulse trains started in phase are presented in Table 2, to a simulation time of 0.172 s. In the first row of each table are presented the results obtained without jitter models.



In third column is shown the total number of coincidences obtained with each model for the simulation time and, in the fourth column is shown the number of coincidence where appears the stop event associate to the best approximation in the frequency measurement process (the desired count is represented like  $10^r$  [7]). Program code for computational experiment in Matlab is presented in [22].

Reference Oscillator	Oscillator under test	Total number of coincidences	Coincidence number for best approximation
Without jitter model	Without jitter model	20223	19999
RJ	RJ	20222	19998
	PJ+RJ	20240	20016
	DCD+RJ	20213	19989
	DDJ+RJ	20223	19999
PJ+RJ	RJ	20217	19993
	PJ+RJ	20224	20002
	DCD+RJ	20218	19995
	DDJ+RJ	20218	19994
DCD+RJ	RJ	20226	20002
	PJ+RJ	20218	19995
	DCD+RJ	20218	19994
	DDJ+RJ	20217	19993
DDJ+RJ	RJ	20227	20003
	PJ+RJ	20216	19991
	DCD+RJ	20218	19994
	DDJ+RJ	20223	20000

Table 2. Simulation results with same magnitude of the jitter and pulse trains in phase.

Comparing the results shown in both columns with the values obtained in the simulation that does not include any jitter model, we see that in the presence of jitter the number of coincidence and the number of coincidence pulse associated with the best approach varies depending on model type. The latter implies the position at the time of the coincidence pulse is not fixed since it depends on the jitter component is dominant, in all case  $n_0=1701023$  and  $n_x=1000000$ .

Fig. 5 is presented the relative error obtained by simulation around best coincidence, obtained under condition (30) with  $\varepsilon=1 \times 10^{-12}$  s. In this plot we can see that due to jitter effect some theoretically expected marginal coincidences may disappear or non-theoretically expected may appear. However, this phenomenon can not be observed in non-marginal coincidences, for instance the expected coincidence under condition (30) and an appropriated pulse width  $\tau$ .

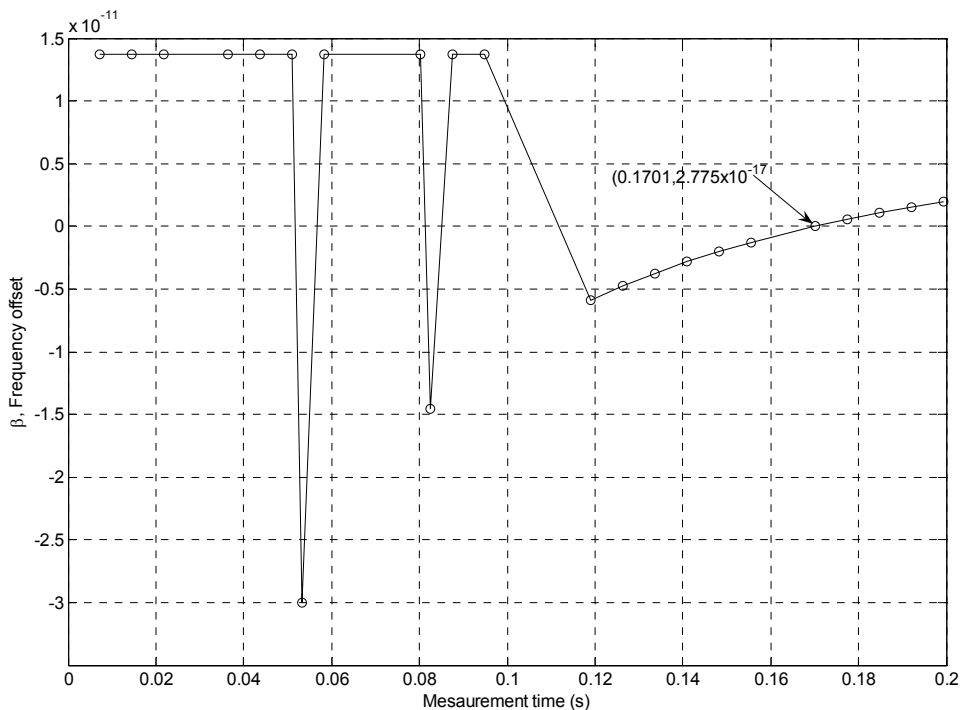


Fig. 5. Jitter effect on frequency relative error for a series of best coincidences (simulation time 0.2.s)

#### 4. Experimental research

For experimental research a prototype of frequency meter was implemented, based on the block diagram shown in Fig 6. This prototype was implement in the FPGA EP3SL150F1152C2N using the development board STRATIX III EP3SL150 and the design software QUARTUS II. In Fig 7. is shown the circuit diagram to the FPGA based frequency measurement prototype used.

For practical frequency measurement, reference frequency was provided by using an Agilent 33250A arbitrary waveform generator and, was set to  $f_0 = 1 \times 10^7$  Hz. The hypothetical unknown frequency was provided by Tektronix AFG3101 Arbitrary function generator.

In Fig. 8a) is presented arbitrary selection from the experimental data set a period value associate to the hypothetical unknown frequency used in experimental research. As shown in Fig. 8a) the measured unknown frequency was adjusted on the rate of  $T_x = 989.97$ ns according to in-built equipment frequency counter. Then this unique pulse train frequency was counted by our experimental prototype of frequency meter. The count of pulses from unknown frequency train obtained with the digital counter Q in block diagram of Fig 6 under the condition (30) is show in Fig. 7 b). The measurement time by single

experimentation never exceeds 1 s. It is evident from screenshot Fig. 7 b) that our prototype circuit has a higher resolution than Tektronix AFG3101 in-built frequency counter, but it shows the same rate of frequency. The character of coincidences between indications of experimental prototype and Tektronix AFG3101 generator was the same on all experimental data set. It permits us to make the next important practical conclusion.

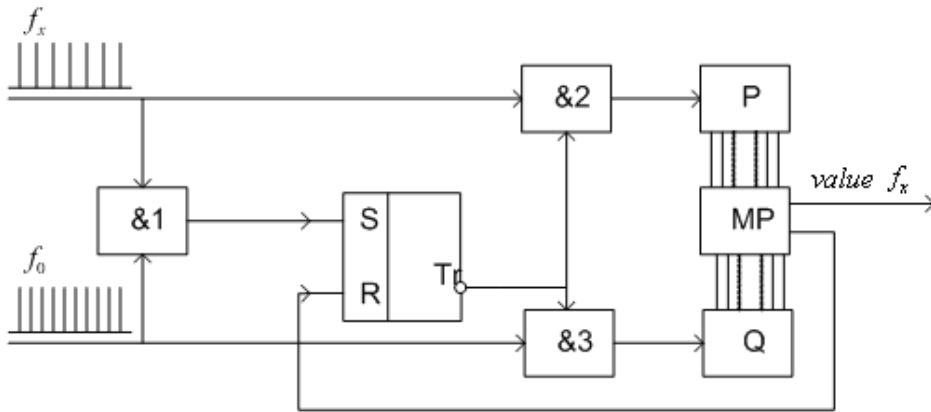


Fig. 6. Frequency meter block diagram.

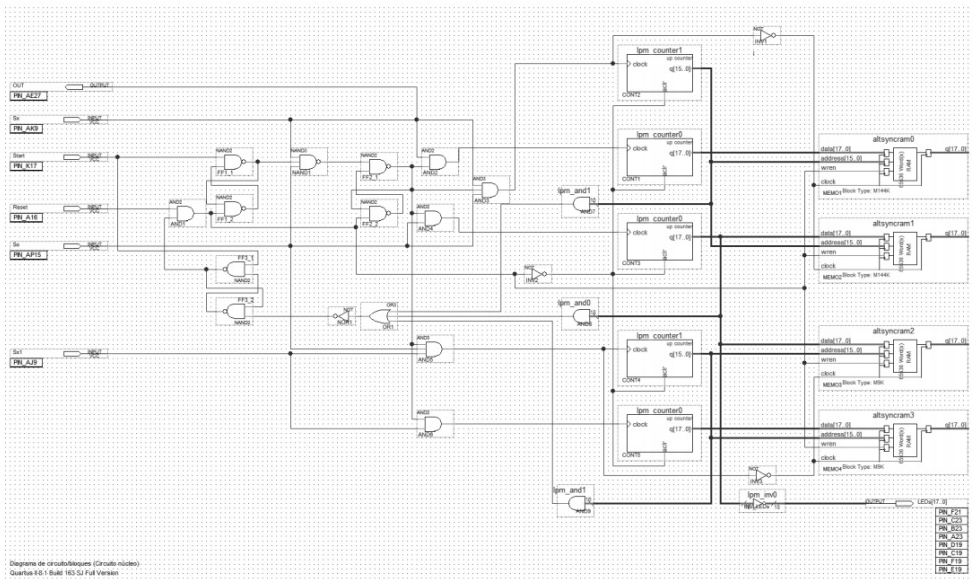
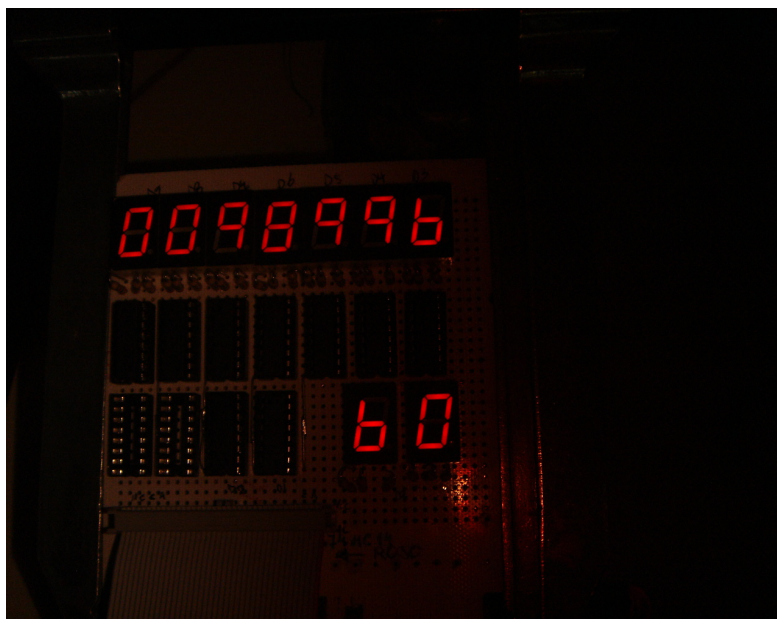


Fig. 7. Circuit diagram for FPGA based frequency meter.



a)



b)

Fig. 8. Sample frequency (formed by arbitrary function generator, a) measurement by prototype circuit (b).

The prototype circuit unfortunately still has multiply sources of noise and uncertainty. But it is a second plan aspect of circuit optimal design in a future work. The main proof of this prototype circuit is the proof of the possibility for a fast frequency measurement without losses of high resolution and rigorous link to actual situation with frequency jitter in a present pulse train.

## 5. Conclusions

In the offered model for fast frequency measurement, the result is fixed on the equality of intervals  $n_0T_0$  and  $n_XT_X$ . Therefore the model is independent to the parameters of coincidence circuits, duration and the shape of coincidence pulses, and the parameters of "zero-crossing" pulses in both sequences.

Due to jitter effect some theoretically expected marginal coincidences may disappear or non-theoretically expected may appear. However, this phenomenon can not be observed in non-marginal coincidences, for instance the expected coincidence under condition (30) and an appropriated pulse width  $\tau$ .

For measuring systems which can be constructed on the basis of the specified model, systematic and instrumental errors have the same infinitesimal order. Instrumental errors are caused only by the reproducibility of the reference frequency.

For measurements of high values of frequency  $f_X$  it is expedient to use higher values of reference frequency in order to have an equivalent reduction of measurement time.

Also it is important to note, that this theoretical method permits to measure unknown frequency value in a case, when unknown frequency exceeds the own value of a standard. For classical methods it is impossible completely.

In future work, there will be an improved version of the experimental prototype to reduce the influence of deterministic errors in measuring and evaluating new frequency estimators based on the Theory of Numbers.

## 6. References

- [1] Stein, S. R. Frequency and Time- Their Measurement and Characterizations. Precision Frequency Control. Vol. 2. Academy Press, New York. 1985. pp 191-416.
- [2] Johansson, S. New frequency counting principle improves resolution. Proceedings of the 36 Annual Precise Time and time interval (PTTI) Systems and Application Meeting, 7-9 December 2004, Washington, D.C., ISBN 0-7803-9053-5, 2005 - 628p.
- [3] Wei Z. The greatest common factor frequency and its application in the accurate measurement of periodic signals. Proceedings of the 1992 IEEE Frequency Control Symposium, pp. 270-273, 1992.
- [4] Fletcher, J. C. Frequency measurement by coincidence detection with standard frequency. U. S. Patent 3, 924,183. 1975.
- [5] Tyrsa V.E. Error reduction in conversion of analog quantities to digitized time intervals. Measurement Techniques. Vol 18, No. 3. 1975. pp. 357-360.
- [6] Tyrsa V.E., Dunashev V.V. Accuracy of frequency measurement base on the pulses coincidence principle. Measurement Techniques. Vol. 24, No.43. pp. 308-312. 1981.

- [7] Hernández-Balbuena D., Sergiyenko O., Tyrsa V., Burtseva, L., Rivas-López M., Signal frequency measurement by rational approximations. Elsevier, "Measurement", Volume 42, Issue 1, Pages 136-144, 2009.
- [8] Hernández-Balbuena D., Oleg Sergiyenko, Vera Tyrsa, Larisa Burtseva *Frequency measurement method for Mechatronic and Telecommunication applications*. IEEE-IES Proceedings of International Symposium on Industrial Electronics (ISIE-2008), Cambridge, United Kingdom, 30 de junio -2 de julio de 2008, 2008, p.1452-1458, ISBN 978-1-4244-1666-0, ISSN/Library of Congress Number 2007936380.
- [9] Hernández-Balbuena D., Oleg Sergiyenko, Vera Tyrsa, Larisa Burtseva. *Method for fast and accurate frequency measurement*. Proceedings of 16th IMEKO TC4 Symposium "Exploring New Frontiers of Instrumentation and Methods for Electrical and Electronic Measurements", Florence, Italy, 20-22 September, 2008. pp. 367-373 [CD-ROM]. ISBN 978-88-903149-3-3.
- [10] Murrieta-Rico, Fabián N., O. Yu. Sergiyenko, V. V. Tyrsa, Daniel Hernandez B. W. ernandez. *Frequency domain automotive sensors: resolution improvement by novel principle of rational approximation*. Proceedings of IEEE-ICIT International Conference on Industrial Technology (ICIT'10), 14-17 March, 2010, Viña-del-Mar, Valparaiso, Chile, pp.1293-1298, ISBN 978-1-4244-5697-0/10.
- [11] Clarkson V, Perkins J. E., and Mareels I. On the novel application of number theoretic methods to radar detection. *Proc. Internat. Conf. Signal Process. Appl. Tech.*, vol. 1, pp. 1202-1211, Oct 1993.
- [12] Clarkson V, Perkins J. E., and Mareels I. Number theoretic solutions to intercept time problems. *IEEE Transactions in Information Theory*, Vol 42. No. 3. pp 959-971, May 1996.
- [13] Tyrsa V., Zenya, A.D. Analysis of errors in frequency comparison by the pulse coincidence method. *Measurement Techniques*, no. 7, pp. 49-51. 1983.
- [14] W. Maichen. *Digital timing Measurement. From scope and probes to timing and jitter*. FRET 33, Frontier in electronic testing. Springer. Netherland. pp 240. 2006. ISBN 0-387-31418-0.
- [15] K. K. Kim et al., Analysis and simulation of jitter sequences for testing serial data channels. *IEEE Transactions on Industrial Informatics*. Vol 4, No 2. pp 134-143. May 2008.
- [16] Buckwalter, et al, Predicting Data-Dependent Jitter. *IEEE Transactions on Circuits and Systems-II:Express briefs*, Vol. 51, No. 9. 2004. pp 543-457.
- [17] B. Analui, et al, Data-Dependent Jitter in Serial Communications. *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, No 11, 2005. pp 3388-3397.
- [18] ITU-T. Vocabulary of digital transmission and multiplexing, and pulse code modulation (PCM) terms. ITU-T Recommendation G.701. International Telecommunication Union. 1994. p 6.
- [19] Sergiyenko O., M. Rivas López, I. Rendón López, V. Tyrsa, L. Burtseva, D. Hernández B. *Possible practical applications of precise optical scanning*. Proceedings of CARS & FOF 07 23rd ISPE International Conference on CAD/CAM Robotics and Factories of the Future, Bogota. Colombia. 16-18 August, 2007, pp. 440-444. ISBN: 978-958-978-597-3.
- [20] Sergiyenko O., W. Hernandez, V. V. Tyrsa, Daniel Hernandez B.. *Precise Optical Scanning for multiuse*. Proceedings of IEEE-35th Annual Conference of IEEE

- Industrial Electronics (IECON'09), 3-5 November, 2009, Porto, Portugal, pp.3399-3404. ISBN 978-1-4244-4649-0/09.
- [21] Rivas-López M., Oleg Sergiyenko, Vera Tyrsa, Wilmar Hernandez Perdomo, Daniel Hernandez B., Luis Devia Cruz, Larisa Burtseva, Juan Iván Nieto Hipólito. *Optoelectronic method for structural health monitoring*. SAGE Publications, International Journal of Structural Health Monitoring, Vol.9, No.2, March, 2010, pp.105-120. Issue Online, September,24, 2009, ISSN 1475-9217 / doi: 10.1177/1475921709340975
- [22] Oleg Segiyenko, Daniel Hernandez B., Vera Tyrsa, Patricia Luz Aurora Rosas Méndez, Moisés Rivas Lopez, Filmar Hernandez, Mikhail Podrygalo, Alexander Gurko. Analisis of jitter influence in fase frequency measurements. Elsevier, "Measurement", Volume 46, Issue 7, Agoust 2011, Pages 1209-1328. ISSN: 0263-2241 - doi:10.1016/j.measurement.2008.04.009

# **Section 4**

## **Optics**





# Optical Radiation Measurement

A. A. Gaertner

*National Research Council of Canada, Ottawa  
Canada*

## 1. Introduction

Optical radiation bathes the world in which we live. The sun, moon and stars have provided us with our primary sources of light since before the beginning of life upon the earth. Very early in our history we observed that we could produce light by other means than depending upon the daily rotation of the earth. As our societies developed, in addition to creating intricate visual experiences with light, we discovered new ways of using this light to effect changes upon many of the materials we found in the world around us. As we sought to control these effects, we realised that the light itself had many properties that affected the outcome of our investigations and that we needed to untangle the characteristics of the light and the properties of the materials in response to this light. We discovered that there were effects caused by 'light' that we could not observe visually, which led to the distinction between light and radiation. To evaluate these effects, to discover new effects, and to produce (or reduce) them efficiently, we have come to depend critically upon the accurate measurement of the optical radiation required in our endeavors. In this chapter we will discuss some of the procedures and equipment necessary to obtain accurate measurements of the amount of optical radiation that we have available for our activities.

The measurement of optical radiation presents the practitioner with many opportunities and challenges. These stem from various sources: the properties of the radiation itself (wavelength, polarization), the geometrical configuration of the radiation field (direction, solid angle), the interaction of the radiation field with the detectors, devices and materials under test (responsivity, reflectance, transmittance), and, not in the least, to the vocabulary and definitions in use, locally and internationally, to describe the measurement process and the results of the measurements.

There are two international organizations whose work is particularly important for the fundamentals of measurement and of optical radiation measurement in particular:

1. BIPM, the Bureau International des Poids et Mesures, ([www.bipm.org](http://www.bipm.org)): "The task of the BIPM is to ensure world-wide uniformity of measurements and their traceability to the International System of Units (SI)."
2. CIE, the Commission Internationale de L'Eclairage, ([www.cie.co.at](http://www.cie.co.at)): "The CIE is devoted to worldwide cooperation and the exchange of information on all matters relating to the science and art of light and lighting, colour and vision, photobiology and image technology."

## 2. Measurements, quantities, units and the SI

As our measurements become more precise and more accurate, it becomes more and more important that we define precisely what we are measuring and what we mean by the terms that we are using to describe the measurements. We must also realize that the purpose of all measurements is to convey information—either to ourselves for some later use, or to someone else somewhere in the world. It is therefore critical that the meanings we associate with the terms we use are the same as the meanings used by the rest of the world. There are four documents that are of particular importance in optical radiation measurements:

1. the VIM (JCGM 200:2008), which provides a vocabulary for the basic concepts and terms of metrology,
2. the GUM (JCGM 100:2008), which provides a guide to the expression of uncertainty in measurement,
3. the SI brochure (BIPM, 2006), which provides a description of the International System of Units, and
4. the CIE Vocabulary (CIE S017, 2011), which provides the vocabulary for optical radiation, radiometry, photometry, lighting and colorimetry.

The concept of measurement involves several important aspects. The basic definition (VIM, Section 2.1) gives measurement as “process of experimentally obtaining one or more quantity values that can reasonably be attributed to a quantity”. In addition to the ‘process’ of the measurement, some examples of which we will discuss in later sections of this chapter, there is the important requirement of a purpose or result to our process, and of a certain form to this result. We must consider a quantity and some means of attributing a value to this quantity. The definitions of quantity and quantity values are given (VIM, Sections 1.1 and 1.19) as:

A *quantity* is a “property of a phenomenon, body, or substance, where the property has a magnitude that can be expressed as a number and a reference”. In this chapter we will discuss the measurement of several specific properties of optical radiation that determine the absolute amount of optical radiation that is available. In particular, we will consider the most important geometrical configurations for the amount of optical radiation (flux, intensity, irradiance, exitance, and radiance) and the spectral quantities associated with each of these geometrical quantities. We will also consider the quantities of optical radiation that are used to describe and measure the response of the human visual system to optical radiation (photometry, colorimetry). These latter quantities are the basis for the design and construction of lighting systems and facilities used for most human endeavors.

A *quantity value* is a “number and reference together expressing magnitude of a quantity”.

In addition to stating that a quantity is something that must have a value, it states the form of a value as containing two parts: a number, and a reference. In all the examples we will be considering, a reference will be an entity called a *measurement unit*, which is defined (VIM, Section 1.9) as a “real scalar quantity, defined and adopted by convention, with which any other quantity of the same kind can be compared to express the ratio of the two quantities as a number”. It can be seen from this definition that measurement units are the basis of our measurements. They are a quantity, just like the one that we are trying to measure (‘same kind’), and they are the particular quantity that we use as the basis to obtain the magnitude

of the quantity we are trying to measure. For example, the quantity of optical radiation flux is measured in terms of a specific amount of optical radiation flux—the watt.

Note that these definitions of quantity and measurement unit contain one of the key parts of a measurement process, which is to compare the two quantities, the unit and the quantity we wish to measure.

The ‘defined and adopted by convention’ aspect of the definition of measurement unit leads us to the SI (BIPM, 2006). This reference quantity (‘real scalar quantity...of the same kind’) could be a quantity of any size that is convenient for our immediate use. However, if the measurement is to convey any information to anyone else—our clients, an industrial factory, another scientist—or even to ourselves at a later time, we need to put qualifications on the properties of this reference quantity. As a minimum, we can see that it needs to be stable over time, and it needs to be accessible to any user of our data. The other users need to either have access to the same unit that we used, or they need to be able to create (realize) one of their own. For true universality, all measurements will need to be based upon reference quantities that have been agreed upon internationally. The present basis of international agreement on units of measurement is the Convention of the Metre (Convention du Mètre), which “is a treaty that created the International Bureau of Weights and Measures (BIPM), an intergovernmental organization under the authority of the General Conference on Weights and Measures (CGPM) and the supervision of the International Committee for Weights and Measures (CIPM)” (<http://www.bipm.org/en/convention/>). The BIPM presently has fifty-five member states, including all the major industrialized countries. The present agreements upon the basic quantities and their units are contained in the SI brochure (BIPM, 2006). This agreement also contains information on the general consensus on how unit symbols and names, quantity symbols, and quantity values should be expressed.

While all these definitions may seem pedantic, they clarify many concepts that we often take for granted without realizing. The length of a piece of wood seems rather straightforward: we have the quantity length (the property of the piece of wood), with its magnitude as a number, such as 2, and a reference (measurement unit), such as the metre. In optical radiation we get into problems quickly because there are so many measurements, quantities and references (measurement units), that it is often difficult to sort out which of the words we are using applies to which of these concepts. The definitions in the VIM and in the SI help us to understand the framework of the measurement process. The definitions in the CIE vocabulary help us to standardize the terms and concepts in optical radiation measurements.

### **3. Traceability chains and measurement uncertainties**

The SI definitions of the basic reference quantities and their units provide the most accurate basis for measurement that is presently available. Although the definitions of these units are constructed such that the units may be realized in any laboratory with the required skills and facilities, the realization of any of these units is not necessarily a simple matter. In general, the primary realization is performed in dedicated national facilities, called National Measurement Institutes (NMIs). At the NMI, this primary measurement standard is established using a measurement procedure that obtains the measurement result, not by

comparison with a reference quantity of the same kind, but by a procedure where the reference is the definition of the measurement unit through its practical realization. In this manner, that primary measurement standard can be said to be “traceable to the SI”, rather than to any other laboratory or NMI. For example, the SI unit for luminous intensity, the candela, is defined in terms of a specific amount of radiant intensity (Section 4.2.2). Calibration of sources based upon the realization of the candela is discussed in Section 7.0.

In all these measurements, the establishment of the reference and our comparison to the reference will be subject to some errors. It is not possible to establish a reference exactly to definition, as it is not possible to compare two quantities exactly. These errors may be systematic or random, leading to a measurement uncertainty, which is defined (VIM, Section 2.26) as a “non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used”. This definition leaves it open as to the parameter used to describe the dispersion, as well as noting that the value depends upon the information used, implying that our information may not be complete and may include an element of belief rather than absolute knowledge (JCGM 104:2009). As a consequence, a measurement result becomes (VIM, Section 2.9) a “set of quantity values being attributed to a measurand together with any other available relevant information.”

Since most measurements are carried out using references that are quantities of the same kind, and are not a primary measurement standard, these measurements and references must somehow be related to the primary measurement standard. This leads to a sequence of calibrations (calibration hierarchy, VIM Section 2.40) that leads from a reference to the final measuring system, where each calibration depends upon the outcome of the previous calibration. The metrological traceability chain is the sequence of measurement standards and calibrations that were used to relate the measurement result to the reference, and the metrological traceability is the property of the measurement result whereby the result can be related to the reference through a *documented unbroken* chain of calibrations, each contributing to the measurement *uncertainty*. An example of a calibration chain used at an NMI is discussed in Section 7.0.

#### 4. Optical radiation quantities and units

Having set up this rather formal measurement structure, we can now consider in more detail the measurement of several basic optical radiation quantities used to determine absolute amounts of optical radiation. Human beings have been dealing with light ever since ‘the beginning’, so they have had a long time to build up a large vocabulary of terms. To understand these terms, we need to consider two important aspects of quantitative optical radiation measurements: the geometrical configuration of the optical radiation that we wish to measure, and the spectral components of this particular geometrical assembly of radiation. Optical radiation is a restricted wavelength range of the electromagnetic spectrum between X-rays and microwave radio waves, from approximately 1 nm to 1 mm (CIE S017, 2011). In this chapter, the focus of our discussions will be upon electromagnetic radiation with wavelengths from approximately 100 nm to 2500 nm, which is an extension from the visible wavelength range, which is approximately from 360 nm to 830 nm.

The basic quantity for the measurement of the amount of electromagnetic radiation is the energy of that radiation, measured with the SI reference unit the joule. This quantity is most

useful when considering pulses or bursts of radiation that have a time limited duration. In this chapter we will deal with the continuous flow of radiation; hence the (energy per time) quantity flux, or power, will be the appropriate basic quantity for the measurement of the amount of radiation. The SI unit for the quantity flux is the watt (W).

The flux of interest will be the flux that is contained within certain geometrical constraints. There is a set of five geometrical configurations that can be used to form the basis of most radiation measurement quantities: total flux, intensity, radiance, exitance and irradiance. The first four of these quantities are characteristics of radiation emanating from a radiation source, whereas the fifth, irradiance, is characteristic of radiation incident upon a surface.

## 4.1 Geometrical quantities and units

### 4.1.1 Solid angle

The geometrical configurations of intensity and radiance make use of the 3-dimensional concept of a solid angle ( $\Omega$ ), which is analogous to the 2-dimensional plane angle ( $\alpha$ ). In Figure 1,  $l$  is the length of the arc of the circle subtended by the limits of the 2-D object at the center of the circle of radius  $r$ . The definition of the plane angle  $\alpha$  is given by

$$\alpha = \frac{l}{r} \quad (1)$$

The SI unit for the plane angle is the radian (rad). A full circle is an angle of  $2\pi$  rad.

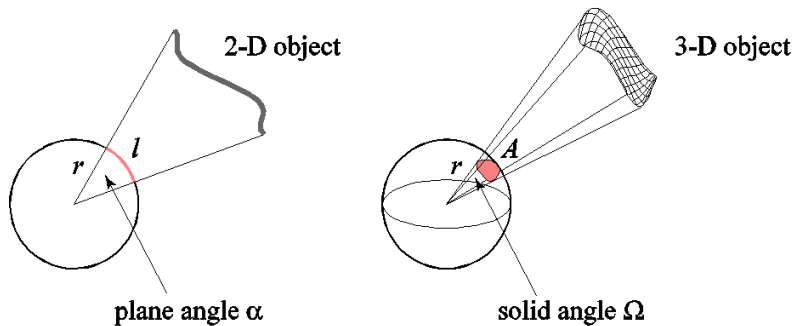


Fig. 1. Two dimensional and three dimensional angles

Similarly, for the 3-D angle, called a solid angle,  $A$  is the area of the part of the sphere of radius  $r$  that is subtended by the limits of the 3-D object at the center of the sphere with radius  $r$ . The definition of the solid angle  $\Omega$  is given by

$$\Omega = \frac{A}{r^2} \quad (2)$$

The SI unit for the solid angle is called the steradian (sr). The complete sphere is a solid angle of  $4\pi$  sr.

### 4.1.2 Total flux ( $\Phi$ )

As indicated above, flux can be measured in many geometrical conditions. However, when radiation sources such as the common incandescent lamps or general lighting service lamps are measured, total flux is used to indicate the total flux output from the lamp into all directions. Specialized equipment, such as integrating spheres or goniophotometers, is required to collect the radiation output into all directions from the lamp (Ohno, 1997). The principles of operation using an integrating sphere are illustrated in Figure 2.

The inside surface of the integrating sphere is coated with a uniformly diffusing material (Lambertian Surface will be discussed in Section 4.1.7.2), such as PTFE (polytetrafluoroethylene) or  $\text{BaSO}_4$ , to provide at the exit port a spatially homogeneous flux that is representative of the total flux output of the lamp. The purpose of the baffle is to prevent radiation directly from the lamp to be incident at the output port, before the sphere has had the opportunity to merge uniformly the radiation output from all directions from the lamp.

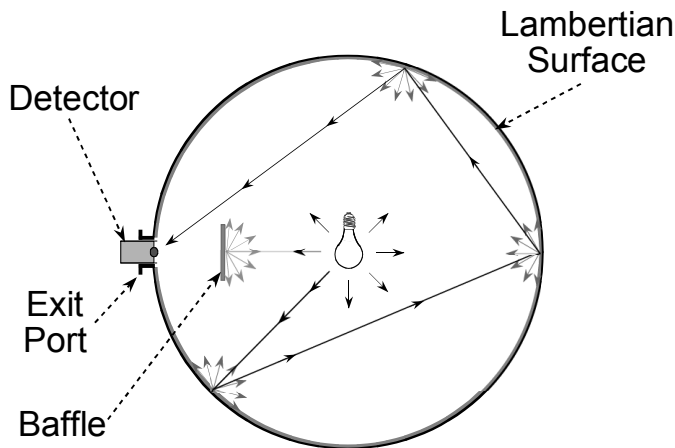


Fig. 2. Total flux measurement using an integrating sphere.

### 4.1.3 Intensity ( $I$ )

Intensity is used rather carelessly in daily use, even in physics. However, in optical radiation measurements it has a very specific meaning (Figure 3). First, the direction from the source ( $\vec{d}$ ) in which the intensity is to be defined must be indicated. Then the intensity of the source in this specified direction is defined as the ratio of the flux ( $\Phi$ ) leaving the source in that particular direction and propagating into an element of solid angle ( $\Omega$ ) containing that specified direction, divided by the size of that element of solid angle. The radiant intensity is defined as

$$I = \frac{\Phi}{\Omega} \quad (3)$$

with SI unit of watt per steradian ( $\text{W sr}^{-1}$ ).

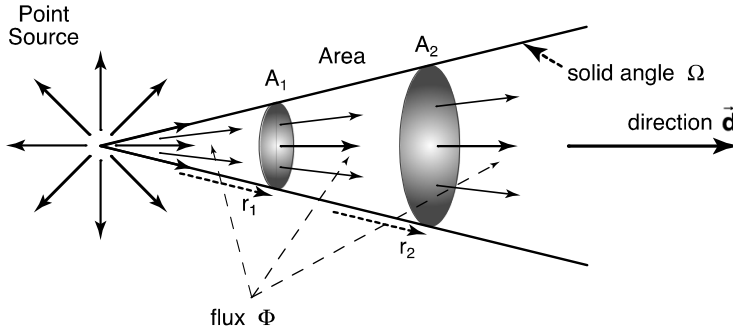


Fig. 3. Defining geometry for the quantity Intensity ( $I$ ) and origins of the inverse square law.

**4.1.4 Irradiance ( $E$ )**

The quantity irradiance is a measure of the amount of radiation incident upon a surface, as illustrated in Figure 4.

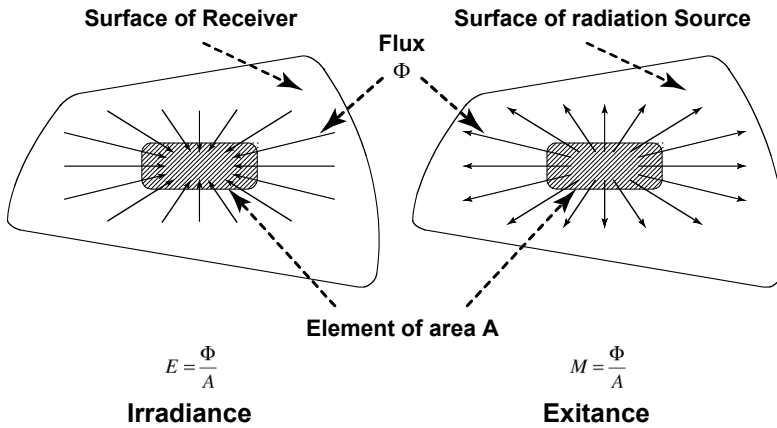


Fig. 4. Defining geometries for the quantities irradiance ( $E$ ) and exitance ( $M$ ).

The irradiance at the surface is defined as:

$$E = \frac{\Phi}{A} \tag{4}$$

where  $\Phi$  is the total flux incident upon the surface element of area  $A$ . Irradiance is measured in terms of the SI unit of watt per square metre ( $W\ m^{-2}$ ).

**4.1.5 Exitance ( $M$ )**

Exitance is very similar to irradiance except that the direction of flow of the flux is reversed—it is leaving the surface, as shown in Figure 4. The radiant exitance from the surface is defined as:



$$M = \frac{\Phi}{A} \quad (5)$$

where  $\Phi$  is the total flux leaving the surface element of area  $A$ . Radiant exitance is measured in terms of the SI unit of watt per square metre ( $\text{W m}^{-2}$ ).

#### 4.1.6 Radiance ( $L$ )

This quantity is the most detailed of the five geometrical quantities we are considering. It takes into account that a radiation source is not a point, but an extended surface, and that a radiation source does not emit the same flux into all directions. Therefore we have a quantity (Figure 5) that specifies the flux ( $\Phi$ ) emitted by a radiation surface from a specified area ( $A$ ) on the surface and in a specified direction ( $\vec{d}$ ) from the surface, and into a specified solid angle ( $\Omega$ ) containing the given direction  $\vec{d}$ . The vector direction  $\vec{n}$  is the perpendicular to the radiation source element of area  $A$ , and  $\alpha$  is the angle between  $\vec{n}$  and  $\vec{d}$ .

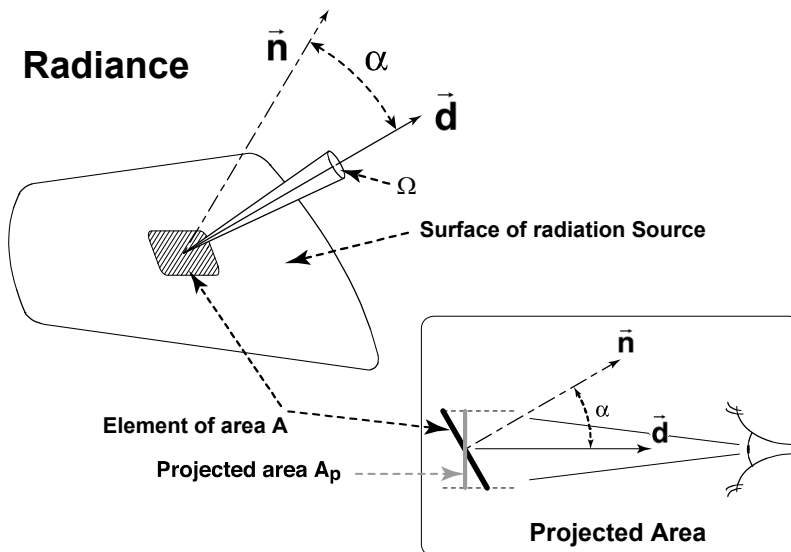


Fig. 5. Defining geometries for the quantity radiance and related projected area.

The radiance of the element of area is defined as

$$L = \frac{\Phi}{A_p \cdot \Omega} \quad (6)$$

The projected area  $A_p$  is the size of the area  $A$  as seen in the specified direction  $\vec{d}$ :

$$A_p = A \cdot \cos \alpha \quad (7)$$

### 4.1.7 Quantity relationships

Several useful relationships between the quantities discussed above may be obtained by adding specific geometric constructions and/or source constraints to the measurement configurations that use these quantities. These are particularly useful to change from a radiation source of one quantity to provide a radiation source of another quantity, such as using an irradiance source to provide a source of radiance.

#### 4.1.7.1 The inverse square law

This is the most common and useful relationship. In its basic form, it relates the irradiance ( $E$ ) of a surface due to the output of a radiant intensity source ( $I$ ) placed at a known distance ( $r$ ) from the surface. The geometry is shown in Figure 3.

The same flux  $\Phi$  passes through each of the areas  $A_1$  and  $A_2$ . Since the sizes of these areas are increasing with distance  $r$  from the source, it is evident that there will be less flux per unit area on surfaces farther from the source. From the definitions of intensity ( $I$ ), solid angle ( $\Omega$ ), and irradiance ( $E$ ), we may derive the relation between  $E$ ,  $I$ , and  $r$ :

$$\Phi = I \cdot \Omega = I \cdot \frac{A}{r^2} \quad (8)$$

The irradiance on any of the surfaces is then obtained as

$$E = \frac{\Phi}{A} = \frac{I}{r^2} \quad (9)$$

This is a very useful concept, but some restrictions should be noted:

1. The source is a point source. A point source is a source that is so small that the inverse square law is valid. If the source is extended, the distance  $r$  from the various points on the source to any area element that is being irradiated will change for all points on the source. The simple equations above are only valid for each point on the source to each point on the irradiated surface. For an extended source, the irradiance will need to be calculated as the sum from all points on the source. An example calculation (Grum & Becherer, 1979) for the case of a circular source of uniform radiance shows the origins of the 'rule of thumb': if the distance  $r$  is at least 10 times larger than the largest dimension of the source or detector, the error will be less than 1%. However, if we depend on changing  $r$  to adjust our required irradiance, good measurement practice would be to test the inverse square behaviour in our actual measurement conditions.
2. The radiation beam considered is not collimated, nor focused by lenses, nor a laser beam. It is the natural propagation of radiant energy from a point on the surface of a radiating source. Any modification of the radiation beam by lenses or the structure of the source will result in the radiant intensity of the source to change with distance from the source, or with the size of the solid angle, or with the direction of propagation with respect to the source. This will invalidate the application of the inverse square law to that particular source. As an example, the measurement of the radiation from LEDs is highly influenced by the geometrical structure in the output of these devices. As a result, the CIE has recommended very specific geometrical conditions for measuring

the intensity of these devices (CIE 127, 2007). The recommended geometries prescribe exact solid angles and distances from the LED, together with specific considerations for the alignment of the LED with the measurement direction. Since the radiant intensity from LEDs is not constant with distance or solid angle, the resulting quantities measured are called Averaged LED Intensities, rather than true intensities as defined in Equation (3) above for a point source.

#### 4.1.7.2 Lambertian surface

A Lambertian Surface is an ideal surface whose radiance is the same in all directions of the hemisphere above the surface. This very special surface has many useful properties. It is basically a radiation source that looks equally bright when viewed from any direction. It can be a radiation source itself, such as a black body radiator, or it can be a material that either transmits or reflects the radiation incident upon it in such a manner. Such a material is also called an isotropic diffusing surface. The basic concepts for such a reflecting diffusing surface are shown in Figure 6.

There are no real materials that have the ideal properties of a Lambertian surface, but pressed PTFE (polytetrafluoroethylene) powder or  $\text{BaSO}_4$  may be used if corrections for their non-ideality are made when required. The reflectance properties of these near-ideal diffusing reflectors are critical to accurate reflectance measurements and their three-dimensional reflection properties are extensively studied (Höpe & Hauer, 2010).

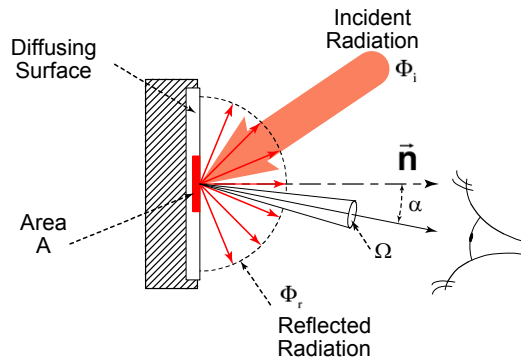


Fig. 6. An isotropic diffusing reflecting surface

#### 4.1.7.3 Irradiance to radiance

The Lambertian diffusing reflecting surface is very useful in obtaining a radiance source from an irradiance source. When determining the relationship between radiometric quantities it is very useful to consider the flow of flux and then add the geometric constraints, as we did in Section 4.1.7.1 above. The geometric configuration shown in Figure 6 shows the conversion from the incident radiation from an irradiance source into the output reflected radiation with the properties of radiance. If the known irradiance at the area  $A$  from our source is  $E$ , then the incident flux  $\Phi_i$  upon the area  $A$  is:

$$\Phi_i = E \cdot A \quad (10)$$

The flux reflected from the area  $A$  is the incident flux times the reflectance ( $\rho$ ) of  $A$ :

$$\Phi_r = \rho \cdot \Phi_i = \rho \cdot E \cdot A \quad (11)$$

We can relate the radiance ( $L$ ) of  $A$  to the reflected flux  $\Phi_r$  using the definition of radiance from Equation (6) and the assumption of a constant radiance in all directions (Lambertian surface):

$$d\phi = L \cdot A \cdot \cos \alpha \cdot d\omega \quad (12)$$

where  $d\phi$  is a small element of flux emitted from  $A$  into the small element of solid angle  $d\omega$ . The total flux emitted from  $A$  by this constant radiance  $L$  is obtained by integrating  $d\phi$  over all the output directions from  $A$ :

$$\Phi_r = \int L \cdot A \cdot \cos \alpha \cdot d\omega = L \cdot A \cdot \int \cos \alpha \cdot d\omega = L \cdot A \cdot \pi \quad (13)$$

Combining Equations (11) and (13) we obtain:

$$L_{IDR} = \frac{\rho \cdot E}{\pi} \quad (14)$$

Equation (14) is the basic equation relating the radiance of an isotropic reflecting diffuser due to a known irradiance of the diffuser, where we have introduced the subscript IDR to indicate that this is the radiance from an Isotropic Diffuse Reflector with reflectance  $\rho$ .

As indicated in the discussion of a Lambertian source, there is no perfect isotropic diffuser. This means that the reflectance  $\rho$  in Equation (14) will not be the same for all directions of both the input irradiation and the output radiance. For accurate conversions from irradiance to radiance we will need to take account of this effect. We introduce a radiance factor ( $\beta$ ) that is defined (CIE S017, 2011) as: *radiance factor* (at a surface element of a non self-radiating medium, in a given direction, under specified conditions of irradiation): ratio of the radiance of the surface element in the given direction to that of the perfect reflecting or transmitting diffuser identically irradiated and viewed. The *reflected radiance factor* is denoted by  $\beta_R$ . For completeness in this description, we also show that the quantities we are discussing are functions of the wavelength  $\lambda$ , as well as angle.

For a perfect reflecting diffuser (PRD) illuminated by an irradiance  $E(\lambda)$ , ( $\rho = 1$ ), so that Equation (14) becomes:

$$L_{PRD}(\lambda) = \frac{E(\lambda)}{\pi} \quad (15)$$

Since this is a perfect reflecting diffuser, this equation is true for all directions of  $L_{PRD}$  from the diffuser.

For an imperfect reflecting diffuser (RD) surface,

$$\beta_R(\lambda, \vartheta) = \frac{L_{RD}(\lambda, \vartheta)}{L_{PRD}(\lambda, \vartheta)} \quad (16)$$

where  $\vartheta$  indicates the direction of the radiance  $L$  from the surface, for some defined irradiation condition. From Equation (16) we obtain:

$$L_{RD}(\lambda, \vartheta) = \beta_R(\lambda, \vartheta) \cdot L_{PRD}(\lambda, \vartheta) = \beta_R(\lambda, \vartheta) \cdot L_{PRD}(\lambda) \quad (17)$$

with the second equality due to the fact that the PRD radiance  $L_{PRD}(\lambda)$  is independent of angle.

Combining Equations (15) and (17) we obtain:

$$L_{RD}(\lambda, \vartheta) = \beta_R(\lambda, \vartheta) \cdot L_{PRD}(\lambda) = \frac{\beta_R(\lambda, \vartheta) \cdot E(\lambda)}{\pi} \quad (18)$$

In our application for converting from irradiance to radiance we usually have the irradiance incident normally upon the diffuser, and the radiance observed at  $45^\circ$ . Therefore, the specific reflectance factor we require is  $\beta_{0/45}(\lambda)$ , where the term 0/45 indicates  $0^\circ$  incidence radiation and  $45^\circ$  output radiance. The final equation relating the input irradiance and the output radiance for this 0/45 geometry is therefore:

$$L_{RD}(\lambda, 45) = \frac{\beta_{0/45}(\lambda) \cdot E(\lambda)}{\pi} \quad (19)$$

The radiance factor  $\beta_{0/45}(\lambda)$  will need to be measured (Höpe & Hauer, 2010) for the particular reflecting diffuser that we are using.

#### 4.1.7.4 Radiance to irradiance

The transfer of radiant energy or flow of radiant flux from a radiating surface to a receiving surface is calculated in a general way for any geometrical configuration using elemental beams of radiance from small areas of the source propagating to small elements of area at the receiver (Grum & Becherer, 1979). The simplest configuration that is often used in calibration laboratories is shown in Figure 7.

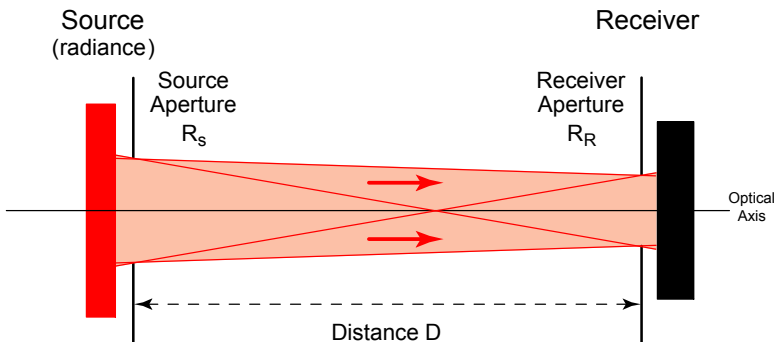


Fig. 7. Irradiance from a Radiance Source

The source and receiver apertures are parallel and centered upon, and perpendicular to, the optical axis. The source is again assumed to be a Lambertian source with radiance  $L$ . The apertures are circular with radii  $R_s$  and  $R_R$  for the source and receiver respectively. Under

these assumptions, the spatial fraction  $f$  of the total flux emitted by the source radiance aperture that is collected within the receiver aperture is given (Walsh, 1958) by:

$$f = 2R_R^2 \left[ (R_S^2 + R_R^2 + D^2) + \sqrt{(R_S^2 + R_R^2 + D^2)^2 - 4R_S^2 R_R^2} \right]^{-1} \quad (20)$$

The radiant exitance of a Lambertian source is obtained by the integration of the constant radiance  $L$  over the hemispherical solid angle to give (Grum & Becherer, 1979):

$$M = \pi \cdot L \quad (21)$$

Therefore the total flux emitted by the source aperture is:

$$\Phi_S = (\pi \cdot L) \cdot (\pi \cdot R_S^2) \quad (22)$$

and the flux collected within the receiver aperture is:

$$\Phi_R = f \cdot \Phi_S = \pi^2 \cdot R_S^2 \cdot f \cdot L \quad (23)$$

The irradiance at the receiver aperture is therefore:

$$E = \frac{\Phi_R}{\pi \cdot R_R^2} = \pi \cdot \left( \frac{R_S}{R_R} \right)^2 \cdot f \cdot L \quad (24)$$

As we will discuss in Section 7, Equation (23) is the basis for the calibration of blackbody sources, which are then used as absolute radiance sources for the realization of spectral irradiance scales using Equation (24).

## 4.2 Spectral quantities and units

The field of optical radiation measurements is often divided into three major categories: radiometry, photometry and colorimetry. Our discussion to this point has been primarily centered upon radiometry, the measurement of the quantities associated with optical radiation. However, human beings have been observing optical radiation for quite some time with their built-in detector—the eye. This has resulted in a very large and important field of radiation measurements. The difference between the three categories is basically that between the characterisation of the electromagnetic radiation that is present in our measurements (radiometry, which is essentially a detector that responds to all wavelengths equally), and the evaluation based on what human beings think is present using their visual detection system composed of the eye and the brain (photometry, colorimetry). Photometry evaluates the radiation as a measure of the strength or magnitude of the human visual response, whereas colorimetry characterizes optical radiation in terms of the human ability to distinguish radiation of different wavelengths, which we call colours.

### 4.2.1 Spectral concepts

The emission, reflection and absorption properties of the sources, detectors and other materials used for optical radiation measurement are all dependent upon the wavelength  $\lambda$

of the radiation under consideration. All the geometrical quantities considered above should be considered to apply at each wavelength we use. The use of the adjective spectral, when applied to any of the quantities  $X$ , indicates (CIE S017, 2011) that either:

1.  $X$  is a function of the wavelength  $\lambda$ , with symbol  $X(\lambda)$ , or
2. the quantity referred to is the spectral concentration, or spectral distribution, of  $X$ , with symbol  $X_\lambda$ . This spectral distribution  $X_\lambda$  is defined as the quotient of the radiant quantity  $dX(\lambda)$  contained in an elementary range of  $d\lambda$  of wavelength at the wavelength  $\lambda$ , by that range:

$$X_\lambda = \frac{dX(\lambda)}{d\lambda} \quad (25)$$

It should be noted that  $X_\lambda$  is also a function of wavelength  $\lambda$ , and to stress this fact, may be written as  $X_\lambda(\lambda)$  without any change in meaning.

As examples, the quantities reflectance ( $\rho$ ) and radiance factor ( $\beta$ ) that we introduced above are functions of wavelength, but not spectral distributions of wavelength. The quantities that involve radiant flux are all spectral distributions. For example, spectral irradiance ( $E_\lambda$ ) is defined (CIE S017, 2011) as the quotient of the radiant power  $d\Phi(\lambda)$  in a wavelength interval  $d\lambda$  incident on an element of a surface, by the area  $dA$  of that element and by the wavelength interval  $d\lambda$ :

$$E_\lambda = \frac{dE(\lambda)}{d\lambda} = \frac{d^2\Phi(\lambda)}{dA \cdot d\lambda} \quad (26)$$

The basic concepts are shown in Figure 8. The SI unit for spectral irradiance is  $\text{W m}^{-2} \text{m}^{-1}$ , or  $\text{W m}^{-3}$ . The unit  $\text{W m}^{-2} \text{nm}^{-1}$  is also used since the wavelength of optical radiation is conveniently measured with the unit nm.

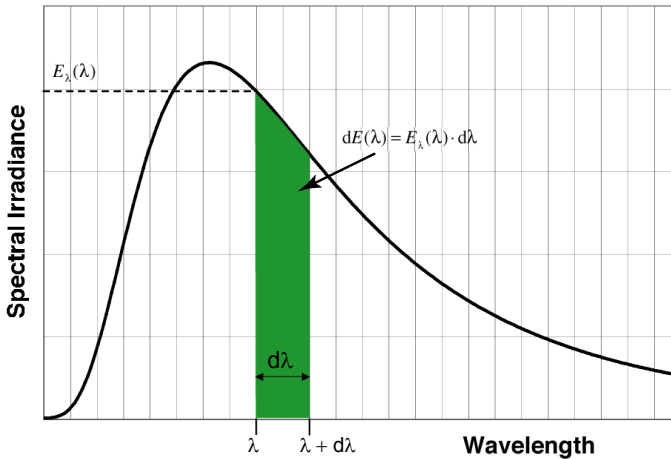


Fig. 8. Spectral distribution components for spectral irradiance

### 4.2.2 Photometry

The retina of the human eye contains two different types of photoreceptors called rods and cones that produce nerve impulses that are passed on to subsequent stages of the human visual system for processing (Ohta & Robertson, 2005). The cones are spread over the entire retina, together with a large concentration within a small central area of our vision called the fovea, which results our high visual acuity at the center of the field of view of the eye. The cones are responsible for our daytime colour vision. The rods are spread over the entire retina except the fovea and are responsible for our night-time, basically black-and-white, vision.

The eye is sensitive to radiation over a range of approximately 11 orders of magnitude from bright sunlight to a flash of light containing only a few photons. The change in size of the pupil area is only capable of controlling the radiation input to the retina by a factor of 12. The remaining adaptation is provided by the rods and the cones. The high radiation range is mediated by the cones and is known as the photopic range. The low radiation range is mediated by the rods and is known as the scotopic range. The intermediate range is mediated by both the rods and the cones, and is known as the mesopic range.

Each human being will perceive differently the amount and colour of a given beam of radiation. In order to simplify calculations and to provide international standards for the measurement of quantities representing the strength of the human visual response to optical radiation, the CIE has standardized spectral weighting functions to be used for each of the three ranges of photometric measurements. The spectral weighting functions for the photopic and scotopic ranges (Figure 9) are known as the spectral luminous efficiency functions for photopic and scotopic vision, with symbols  $V(\lambda)$  and  $V'(\lambda)$  respectively (CIE Standard S010/E:2004). The mesopic range requires a more complex weighting function based upon a gradual transition between  $V(\lambda)$  and  $V'(\lambda)$  throughout the mesopic region that depends on the visual adaptation conditions (CIE 191:2010).

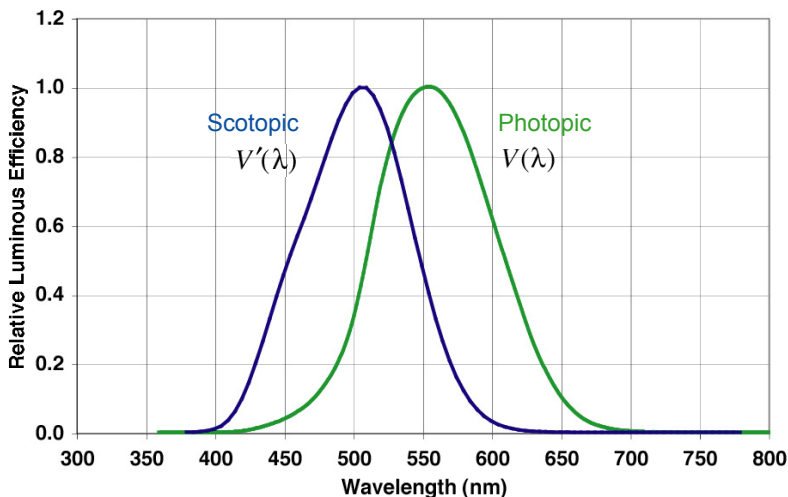


Fig. 9. CIE spectral luminous efficiency functions



The CIE spectral luminous efficiency functions are relative spectral values with a maximum value of one. Within the SI, photometric quantities and units are obtained by calculation from radiometric quantities and units. This is facilitated with the definition of a quotient  $K$  called the luminous efficacy of radiation and defined as (CIE S017, 2011) the quotient of the luminous flux,  $\Phi_v$ , by the corresponding radiant flux  $\Phi_e$ . (The subscripts v and e are used to indicated photometric and radiometric quantities respectively.)

$$K = \frac{\Phi_v}{\Phi_e} \quad (27)$$

The spectral luminous efficiency functions provide the relative spectral values for  $K(\lambda)$ . In order to obtain the absolute scaling, there needs to be a defined scaling factor between radiometry and photometry at one of the wavelengths of the luminous efficiency functions. This is done in the definition (BIPM, 2006) of the SI base unit of luminous intensity, the candela:

“The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency  $540 \times 10^{12}$  hertz and that has a radiant intensity in that direction of  $1/683$  watt per steradian. It follows that the spectral luminous efficacy for monochromatic radiation of frequency of  $540 \times 10^{12}$  hertz is exactly 683 lumens per watt,  $K = 683 \text{ lm/W} = 683 \text{ cd sr/W}$ .”

This value of  $K$  is the scaling factor for all the luminous efficiency functions. The value 683 was chosen to provide the best continuity between the size of the candela historically established using candles and platinum blackbodies before the change in 1979 to link the photometric units to the radiometric units. The definition is given for a specific frequency of radiation, which corresponds to approximately 555 nm in air (555.016 nm in standard air). The definition in terms of frequency avoids wavelength problems caused by the different index of refraction of different media of propagation. This does complicate slightly the determination of the maximum luminous efficacy value  $K_m$  that is used to obtain photometric quantities from radiometric quantities using the CIE spectral luminous efficiency functions that are given as functions of wavelength (CIE S017, 2011). From these definitions we have (CIE S017, 2011): For photopic vision  $K_m = 683 V(555 \text{ nm})/V(555.016 \text{ nm}) \text{ lm} \cdot \text{W}^{-1} = 683.002 \text{ lm} \cdot \text{W}^{-1} \approx 683 \text{ lm} \cdot \text{W}^{-1}$ . For scotopic vision  $K'_m = 683 V'(507 \text{ nm})/V'(555.016 \text{ nm}) \text{ lm} \cdot \text{W}^{-1} = 1700.05 \text{ lm} \cdot \text{W}^{-1} \approx 1700 \text{ lm} \cdot \text{W}^{-1}$ . For all other wavelengths  $K(\lambda) = K_m V(\lambda)$  and  $K'(\lambda) = K'_m V'(\lambda)$ .

From these definitions, all the geometrical photometric quantities may be determined from the corresponding radiometric quantities discussed in Section 4.1 above. For example,

$$\begin{aligned} \Phi_v &= K_m \int \Phi_e(\lambda) \cdot V(\lambda) \cdot d\lambda \\ \Phi'_v &= K'_m \int \Phi_e(\lambda) \cdot V'(\lambda) \cdot d\lambda \end{aligned} \quad (28)$$

where  $\Phi_v$  is the photopic luminous flux in units of lumens,  $\Phi'_v$  is the scotopic luminous flux, also in units of lumens,  $\Phi_e(\lambda)$  is the spectral distribution of radiant flux in units of watts per nanometre and  $d\lambda$  is the wavelength in units of nanometre.

The relationships between the radiometric and photometric quantities, units and symbols are shown in Table 1. Although luminous exitance has the same unit (lumen per square metre) as illuminance, the unit lux is reserved for illuminance only.

Radiometry		Photometry	
<i>Quantity (symbol)</i>	<i>Unit (symbol)</i>	<i>Quantity (symbol)</i>	<i>Unit (symbol)</i>
Radiant Energy ( $Q_e$ )	joule (J)	Luminous Energy ( $Q_v$ )	lumen second (lm s)
Radiant Flux ( $\Phi_e$ )	watt (W)	Luminous Flux ( $\Phi_v$ )	lumen (lm)
Radiant Intensity ( $I_e$ )	watt per steradian (W/sr)	Luminous Intensity ( $I_v$ )	candela (cd)
Radiance ( $L_e$ )	watt per steradian per square metre (W/(sr m <sup>2</sup> ))	Luminance ( $L_v$ )	candela per square metre (cd/m <sup>2</sup> )
Irradiance ( $E_e$ )	watt per square metre (W/m <sup>2</sup> )	Illuminance ( $E_v$ )	lumen per square metre (lm/m <sup>2</sup> ) lux (lx)
Radiant Exitance ( $M_e$ )	watt per square metre (W/m <sup>2</sup> )	Luminous Exitance ( $M_v$ )	lumen per square metre (lm/m <sup>2</sup> )

Table 1. Radiometric and Photometric quantities, units and symbols.

#### 4.2.3 Colorimetry (Ohta & Robertson, 2005)

The ability of humans to distinguish radiation of different colours is due to the three types of cone cells that are present in the retina (Ohta & Robertson, 2005). Based upon measurements of the human visual response, the CIE has standardized a system of colorimetry (CIE 015:2004) that enables color to be specified precisely for any arbitrary color stimulus. This XYZ color specification system is based upon the calculation of three colorimetric CIE tristimulus values  $X$ ,  $Y$  and  $Z$  of a radiation source or colour stimulus  $S(\lambda)$  as given in equation (29).

$$\begin{aligned}
 X &= K_m \int S(\lambda) \bar{x}(\lambda) d\lambda \\
 Y &= K_m \int S(\lambda) \bar{y}(\lambda) d\lambda \\
 Z &= K_m \int S(\lambda) \bar{z}(\lambda) d\lambda
 \end{aligned} \tag{29}$$

The spectral shapes of the three CIE color matching functions  $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$  and  $\bar{z}(\lambda)$  are shown in Figure 10. A virtual observer having these color matching functions is called the CIE 1931 Standard Colorimetric Observer or the CIE 2° Colorimetric Observer, where the 2° indicates that the color matching functions are based upon those measured for a visual 2° field of view. Similar values are available for a 10° field of view, called the CIE 1964 Standard Colorimetric Observer or the CIE 10° Colorimetric Observer.

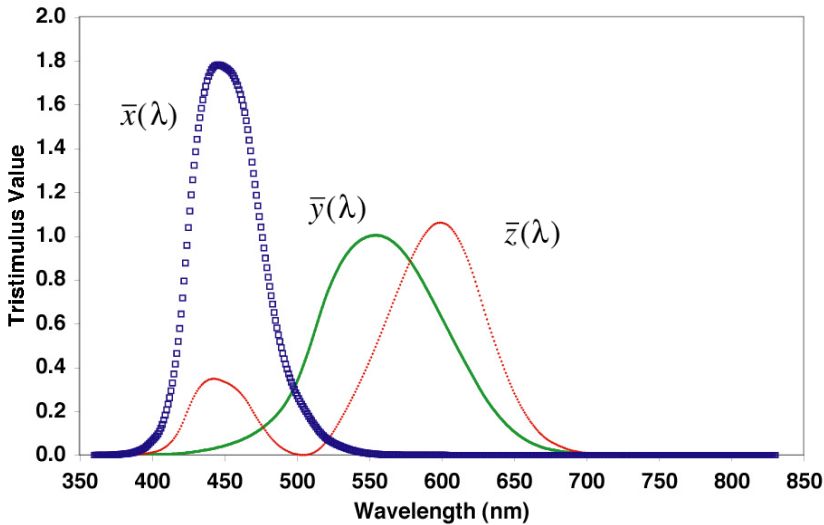


Fig. 10. CIE color matching functions

#### 4.2.4 Correlated Colour Temperature (CCT)

Many of the radiation sources used as standards for radiometric and photometric measurements are incandescent lamps whose relative spectral output closely resembles that of a Planckian radiator or blackbody. Since the spectral radiation from a Planckian radiator can be described with the single temperature variable, it is often convenient to describe the colour of an incandescent source using only one variable similar to the absolute temperature used to describe the radiation from a Planckian radiator. For example, the electrical operating variables (current, voltage) of incandescent lamps used for many photometric standards and measurements are set such that the correlated colour temperature of the radiation from the lamp is approximately 2856 K. This provides a spectral distribution that is reasonably reproducible for use in many photometric and colorimetric applications. The CIE Standard Illuminant A is a specified relative spectral distribution equal to that of a blackbody at a temperature of approximately 2856 K (CIE S 014-2, 2006)

The definition of correlated colour temperature (CCT) is standardized by the CIE (CIE S017, 2011) such that it may be calculated from the measured spectral radiation. The complete colour specification of a radiation source or color stimulus  $S(\lambda)$  is three colorimetric tristimulus values as given in Section 4.2.3 above. In addition, the CIE defined the  $\bar{y}(\lambda)$  colour matching function to be equal to the photometric  $V(\lambda)$  spectral weighting function, so that the Y tristimulus value holds all the information about the source brightness. This enabled the definition (Equation 30) of two normalized quantities  $x$  and  $y$ , called chromaticity coordinates, that contain the colour information about the source.

$$\begin{aligned} x &= X/(X + Y + Z) \\ y &= Y/(X + Y + Z) \end{aligned} \quad (30)$$

This enables the representation of colours in two-dimensional plots. Note that for determination of the chromaticity coordinates, it is only necessary to know the relative spectral distribution for  $S(\lambda)$ . The chromaticity coordinates are shown in Figure 11 for several colour stimuli. The bounding curve for color stimuli is given by the spectrum locus, which is composed of the chromaticity coordinates of all the pure monochromatic wavelengths of radiation, and the purple boundary, which is the line joining the ends of the spectrum locus. The chromaticity coordinates of blackbody radiators form a smooth curve within the spectrum locus. This smooth curve gives us the basis for our single variable color temperature definitions. If the source has a relative spectral distribution equal to a Planckian radiator, its chromaticity coordinates will fall upon this curve and the source is said to have a color temperature  $T_c$  equal to that of the corresponding blackbody radiator with the same chromaticity coordinates. The radiation from most practical sources will have a chromaticity that does not equal that of a blackbody. In this case their correlated colour temperature  $T_{cp}$  is defined as the temperature of the blackbody whose chromaticity is nearest to that of the radiation. There are several methods of determining  $T_{cp}$  from the relative spectral distribution of the stimulus  $S(\lambda)$  (Ohta & Robertson, 2005, Gardner, 2000).

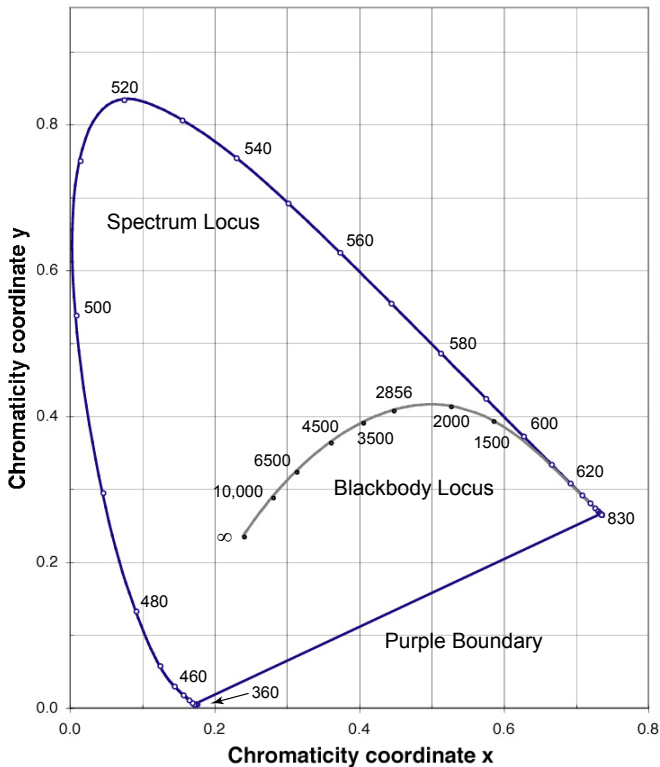


Fig. 11. CIE xy chromaticity diagram of the XYZ color specification system. The points indicated along the Spectrum Locus are monochromatic wavelengths in units of nanometres. The points indicated along the Blackbody Locus are temperatures of Planckian radiators in units of Kelvin.

#### 4.2.5 Spectral mismatch correction factor ( $F^s$ )

The definitions of photometric and colorimetric quantities given above require knowledge of the spectral distribution of the radiation that we are measuring. Most of the instruments in daily use for photometric and colorimetric measurements are simple devices that do not contain the spectroradiometers necessary for the measurement of spectral distributions. The simplest photometers and colorimeters are constructed from detectors, usually of silicon, and transmitting color glass filters. The types and thicknesses of the 3-to-4 types of glasses are chosen such that the relative spectral responsivity of the combination of the detector plus filter is as close to the desired photometric or colorimetric functions (e.g.  $V(\lambda)$ ,  $V'(\lambda)$ ,  $\bar{x}(\lambda)$ ,  $\bar{y}(\lambda)$  or  $\bar{z}(\lambda)$ ) as possible for the construction cost involved. These devices produce a single-number result for the measurement. Since the output of a detector is a voltage or current, this value must be converted to the corresponding photometric or colorimetric quantity by a calibration of the device, which is usually performed by a measurement of a known amount of the desired quantity. This produces a calibration factor that is often built into the electronics of the device.

These CIE spectral functions are not easy to reproduce in this physical form and there will always be errors or uncertainties in measurement using these devices. An estimate of the error involved can be determined by calculating what is called a Spectral Mismatch Correction Factor ( $F^s$ ). To derive this factor we compare the quantity that we wish to measure with the quantity that we are actually measuring, using illuminance as an example as shown in Figure 12. The photometer in this case is called an illuminance meter or luxmeter.

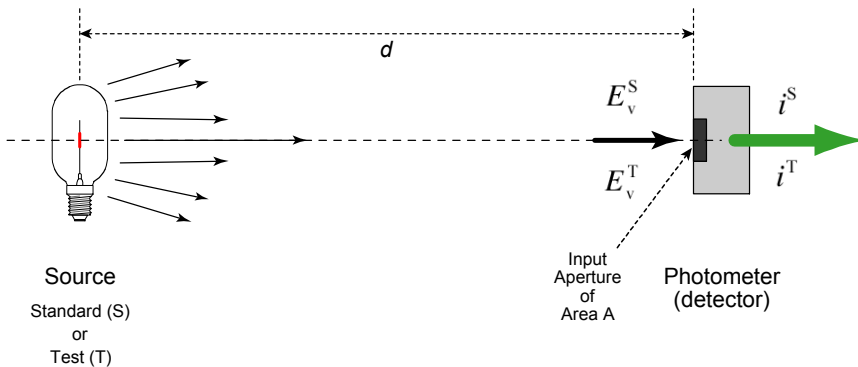


Fig. 12. Basic measurement and calibration configuration

The variables that describe the measurement shown in Figure 12 are:

Required value:	$E_v^T$ , the illuminance produced by the test source.
Known value:	$E_v^S$ , the illuminance produced by the standard source.
Measured values:	$i^S$ , the photometer output for the standard source.
	$i^T$ , the photometer output for the test source.

Note that the output of the photometer is a single number.

We use the material properties of the source output and the photometer components to determine the origin of our measured photometer outputs. The variables are:

$P_e^T(\lambda)$ , the relative spectral irradiance produced by the test source,  
 $P_e^S(\lambda)$ , the relative spectral irradiance produced by the standard source, and  
 $R(\lambda)$ , the relative spectral responsivity of the photometer.

The geometrical factors that convert these relative values to absolute values may be written as  $a$ ,  $b$  and  $r$  respectively. The photometer outputs for the two sources can then be written:

$$\begin{aligned} i^T &= a \cdot r \cdot \int_{\text{all wavelengths}} P_e^T(\lambda) \cdot R(\lambda) \cdot d\lambda \\ i^S &= b \cdot r \cdot \int_{\text{all wavelengths}} P_e^S(\lambda) \cdot R(\lambda) \cdot d\lambda \end{aligned} \quad (31)$$

The measurements we require are the two illuminances:

$$\begin{aligned} E_v^T &= a \cdot v \cdot \int_{360 \text{ nm}}^{830 \text{ nm}} P_e^T(\lambda) \cdot V(\lambda) \cdot d\lambda \\ E_v^S &= b \cdot v \cdot \int_{360 \text{ nm}}^{830 \text{ nm}} P_e^S(\lambda) \cdot V(\lambda) \cdot d\lambda \end{aligned} \quad (32)$$

Note that  $V(\lambda)$  is exactly zero outside the 360-830 nm range, whereas the sources and the photometer might very well be contributing a signal outside this range. By taking ratios of the equations we may eliminate the geometrical constants and scaling factors  $a$ ,  $b$ ,  $r$  and  $v$  to obtain the desired illuminance of the test source:

$$\begin{aligned} E_v^T &= \frac{E_v^S}{i^S} \cdot i^T \cdot \frac{\int_{360 \text{ nm}}^{830 \text{ nm}} P_e^T(\lambda) \cdot V(\lambda) \cdot d\lambda}{\int_{\text{all wavelengths}} P_e^T(\lambda) \cdot R(\lambda) \cdot d\lambda} \cdot \frac{\int_{\text{all wavelengths}} P_e^S(\lambda) \cdot R(\lambda) \cdot d\lambda}{\int_{360 \text{ nm}}^{830 \text{ nm}} P_e^S(\lambda) \cdot V(\lambda) \cdot d\lambda} \\ &= CF \cdot i^T \cdot F^* \end{aligned} \quad (33)$$

where CF is a photometer Calibration Factor that is stored in the electronics of the photometer when the photometer is calibrated with the standard source. This CF converts the measured  $i^T$  into  $E_v^T$  when the test source is measured. However, the  $F^*$  term is not included in the photometer measurement or calibration. It can only be determined at the point of measuring the test source, since  $P_e^T(\lambda)$  is required. In general,  $P_e^T(\lambda)$ ,  $P_e^S(\lambda)$  and  $R(\lambda)$  are not known for most photometers and measurements. A standard source that approximates CIE Source A is usually used for the calibration of photometers. The CIE functions are defined, and generic values for  $P_e^T(\lambda)$ ,  $P_e^S(\lambda)$  and  $R(\lambda)$  may be used to obtain an estimate of the  $F^*$  for the photometer or colorimeter and radiation test source in use. The  $F^*$  can approach 1.0 under one or both of two conditions: 1) if the relative spectral distributions of the standard source and the test source are the same or proportional, i.e.  $P_e^T(\lambda) \propto P_e^S(\lambda)$ . This equivalence is one of the bases for the incentive in measurement or calibration to compare 'like-with-like'. 2) if the relative spectral responsivity of the photometer is proportional to the CIE function ( $V(\lambda)$  in our example above). This requires some effort on the part of manufacturers and photometers/colorimeters of many different qualities and cost are available. The evaluation of the quality of photometers and colorimeters has been documented by the CIE in various publications (CIE 179:2007, CIE 069:1987).

## 5. Measurements and calibrations

In Section 2 we noted that one of the key parts of a measurement process was to compare two quantities, the quantity we wish to measure and the quantity unit. For most measurements the quantity unit is embodied in a calibrated measurement standard whose quantity value is a multiple or submultiple of the base unit. For example, a standard luminous intensity lamp may produce a luminous intensity of 219.2 cd. To measure the luminous intensity output of another (test) light source, we will need some means of comparing the output luminous intensity of the test source with that of the standard source. This involves the introduction of a third device, that of a transfer device. In our example, this device must be able to compare the luminous intensity of both sources and return values for each lamp that are proportional to the luminous intensity of the lamp, or a value that is the ratio of the luminous intensities of the two sources.

For accurate measurements, it is imperative that the transfer device be capable of comparing the exact quantity that we wish to measure. This includes both the geometrical aspects and the spectral aspects that we discussed in Section 4. The general schematic for this process is shown in Figure 13.

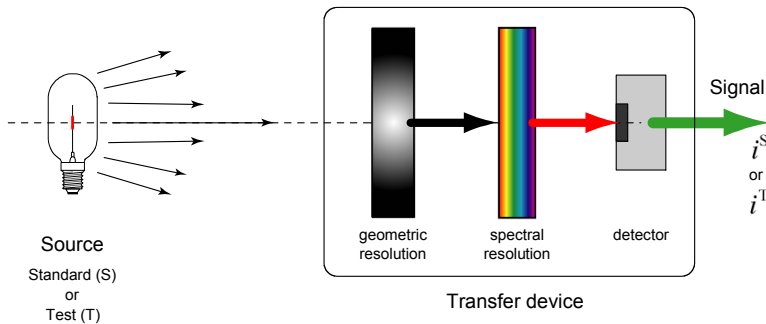


Fig. 13. Measurement Transfer Device

The spectral resolution component indicated for the transfer device is not necessarily the only spectrally important component. In the case of photometric and colorimetric measurements the spectral responsivity of the complete transfer device must be equal to the desired CIE function. If this is not the case, an  $F^*$  will need to be determined as indicated in Section 4.2.5 above.

If the transfer device is only used to measure the output of the test source by comparison with the standard source, the measured value of the test source is simply:

$$S^T = \frac{i^T}{i^S} \cdot S^S = \frac{S^S}{i^S} \cdot i^T \quad (34)$$

where  $S^T$  and  $S^S$  are the desired quantities for the test source and the standard source, and  $i^T$  and  $i^S$  are the signals from the transfer device when measuring these sources. An  $F^*$  is to be applied to Equation (34) as necessary. If the standard source is a measurement

standard with accompanying uncertainties, and the uncertainties of the transfer process are accounted for, the test source is now said to be calibrated. (VIM, Section 2.39).

If the intent is to use the transfer device for subsequent measurements of the same quantity,

the factor  $\frac{S^S}{i^S}$  may be considered a calibration factor for the transfer device, to be used to

convert the subsequent  $i^T$  measurements into the corresponding  $S^T$ . If the uncertainties of the standard and the process of measurement of the standard by the transfer device are considered and included, the transfer device may now be said to be calibrated for the measurement of the quantity of the standard. The calibration factor will have units of the ratio of the units of the standard source to the units of the signal. In the case of an illuminance meter calibration using the configuration of Figure 12, the units of the calibration factor for the photometer would be lux per ampere.

Figure 13 shows the transfer device as a detector unit. It is also possible that a radiation source be a transfer device. For example, if we wish to transfer a calibration between two photometers, from a calibrated standard photometer to a test photometer, we would use an incandescent lamp as the transfer device. Although it would not be necessary that this lamp be calibrated, the lamp should produce a geometric and spectral radiation field that is appropriate for the measurements to be performed by the photometers and the quantity calibration to be transferred between the photometers.

There are many measurement configurations possible that depend upon the type of measurement standard that is available and the required measurement. We will discuss several of the common configurations in the following subsections.

### 5.1 Total flux

The measurement transfer device for total flux is the integrating sphere and detector unit shown in Figure 2, where the detector could be a spectroradiometer for spectral radiant flux measurements. The standard source and the test source are sequentially placed at the center of the sphere and the signal from their flux output is recorded at the detector. The measured/calibrated total flux for the test source is given by Equation (34), with an  $F^*$  applied for photometric measurements. In addition to the detector, the integrating sphere walls and all the interior components such as baffles and lamp supports must be considered as part of the responsivity  $R(\lambda)$  of this transfer device. If the standard source and the test source have different output spatial distributions, any geometrical differences in the responsivity of the sphere may need to be considered. A detailed consideration of integrating sphere measurements has been given by (Ohno, 1997).

### 5.2 Intensity and irradiance/illuminance

These two geometric quantities are often considered together since the measurement configuration, shown in Figure 12, is the same for both. The analysis of the measurement equations is based upon the observation that the output of a radiation detector is directly dependant upon the total flux that is incident upon the sensitive area of the detector, which will be assumed in this chapter to be the area of the input aperture. Therefore, the



basic equation relating the two detector signals and the radiation from the sources is simply:

$$\frac{i^T}{i^S} = \frac{\Phi^T}{\Phi^S} \quad (35)$$

The geometric resolution components placed before the detector, such as shown in Figure 13, are designed to ensure that only the desired flux is input to the detector. In the case of intensity and irradiance/illuminance measurements, the geometrical factors are the distance  $d$  and the aperture area  $A$ . Although the intensity of a (point) source does not change with distance, the flux  $\Phi$  that is incident upon a detector with a fixed input aperture  $A$  varies with the square of the distance as we discussed in Section 4.1.7.1. Therefore any comparison of two intensity sources (or irradiance/illuminance sources) will depend upon the distance at which we measure their output using a detector. If we allow the distances of the test and standard intensity sources from the detector aperture to be different,  $d^T$  and  $d^S$  respectively, the detector signal ratio is given by:

$$\frac{i^T}{i^S} = \frac{\Phi^T}{\Phi^S} = \frac{I^T \cdot A}{(d^T)^2} \cdot \frac{(d^S)^2}{I^S \cdot A} = \frac{I^T}{I^S} \cdot \left( \frac{d^S}{d^T} \right)^2 \quad (36)$$

from which we obtain the test source intensity as

$$I^T = I^S \cdot \left( \frac{d^T}{d^S} \right)^2 \cdot \frac{i^T}{i^S} \quad (37)$$

To reduce errors and uncertainties, we usually try to set up our measurements such that  $d^T \approx d^S$  and that this distance is large enough that the inverse square law is valid.

If we combine Equations (8) and (9) together with Figure 12, we see that an intensity source may be used to calibrate an irradiance/illuminance meter:

$$E^S = \frac{\Phi^S}{A} = \frac{I^S}{(d^S)^2} \quad (38)$$

The input aperture area of the detector has cancelled out of the equations.

### 5.3 Radiance/luminance

A very convenient method used to produce a standard radiance or luminance source uses an isotropic diffusing reflecting surface to produce a known radiance from a known irradiance incident upon the surface, as discussed in Section 4.1.7.3 above. A schematic for the measurements is given in Figure 14. Note that the known irradiance at the diffuser may be produced by either a standard radiant intensity source (Equation (38)) or a standard irradiance source. For luminance, the known sources may also be either a standard luminous intensity source or a standard illuminance source.

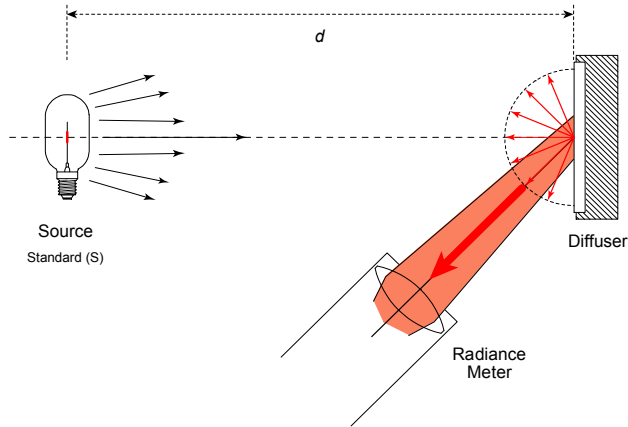


Fig. 14. Radiance source derived from an irradiance source with diffuser

## 5.4 Optical radiation measurement standard sources

The standard radiation sources discussed in the above examples are generally incandescent lamps. Since their output cannot be calculated directly from the properties of the lamps, these sources must be calibrated from primary measurement standards traceable to the SI as indicated in Section 3 above and discussed in Section 7 below. However, with care, incandescent lamps are excellent secondary, reference and working measurement standards that can be used to transfer calibrations from the primary standards realised in NMIs to the working standards used in many calibration laboratories. The choice and use of these lamps as standards requires consideration of their spectral, geometrical, electrical, mechanical and recalibration characteristics (CIE 149:2002).

### 5.4.1 Spectral characteristics

The output spectral distributions of incandescent lamps are particularly suitable for use as standards because the spectral distribution is a smooth and continuous function of the wavelength such as shown in Figure 8. This reduces the errors and uncertainties associated with measurements of their spectral distributions in narrow wavelength bandwidths. There are three main types of incandescent lamps characterized by the different gas fillings. Bulbs with no filling, or vacuum bulbs, can be used up to a maximum CCT of approximately 2400 K before the effects of filament evaporation become unacceptable. General lighting service bulbs filled with nitrogen and an inert gas at approximately atmospheric pressure may be used from approximately 2000 K to 2900 K CCT before the evaporation becomes excessive. Tungsten-halogen lamps, which operate with a gas filling pressure between 7 to 10 times atmospheric pressure and with a hot quartz envelope, may be used up to CCTs of approximately 3400 K.

### 5.4.2 Geometrical characteristics

The construction of the bulb and the filament characteristics of these lamps can be somewhat adjusted to facilitate their application to the different geometrical quantities discussed.

The lamps used for total flux in integrating spheres are usually designed with spherical bulbs and circular or distributed filament shapes to provide a uniform spatial output for a uniform illumination of the sphere walls.

The lamps used for intensity and illuminance are designed with planar filaments that enables the distance between the lamp and detector to be determined reproducibly and accurately. The bulb of the lamp is also shaped, such as a triangular shape, to reduce the inter-reflections inside the bulb that will cause scattered light errors in the measured output of the lamp. Spectral irradiance lamps are usually of the tungsten-halogen design operating with a CCT of approximately 3200 K to increase the amount of UV radiation. The deviation of the irradiance from these lamps from the inverse-square-law should be measured if they are to be used at different distances from that at which they were calibrated.

Sources used for luminance and radiance standards are often the irradiance/diffuser combination described in Section 5.3. If a higher radiance is required, ribbon or strip-filament lamps can be used. Present lamps of this type are limited to approximately 2900 K.

#### **5.4.3 Electrical characteristics**

The filament of the incandescent lamp is heated by passing an electrical current through it. For measurement standard lamps, this is usually a direct current. The same polarity of the electrical current must be used each time the lamp is operated. To avoid thermal shock to the filament, this current should be applied gradually, over times on the order of a minute or more. When the lamp is calibrated, the current or voltage is adjusted until the spectral radiant output of the lamp reaches the desired operating CCT. This electrical operating point, either a defined operating current or voltage, must be applied whenever the lamp is used to obtain the same radiant output each time the lamp is used. Since all incandescent lamps will age with use, the second of the two electrical quantities (voltage, if the lamp current is the defined operating variable) may be used as a monitor of the ageing of the filament and consequent degradation of the lamp output calibration.

The electrical quantities must be measured accurately since the radiant output of an incandescent lamp depends strongly upon the electrical power applied to the lamp. Since the electrical power causes a change in temperature, the spectral distribution of the radiant output changes as well as the absolute amount of output. For lamps operating with a CCT of approximately 2856 K, it has been observed (CIE 149:2002) that the luminous output of the lamp changes approximately 4% for a 1% change in lamp operating voltage and approximately 8% for a 1% change in lamp operating current.

#### **5.4.4 Mechanical characteristics**

Incandescent lamps are sensitive to vibration and shock. In addition to breakage of the glass envelope or the electrical feed-throughs, the filament structure is particularly fragile. As indicated above, electrical power should be applied gradually to avoid thermal shock.

The electrical and mechanical properties of a lamp change rapidly when a lamp is first used. Before a lamp is calibrated, it is usually aged by operating at the desired CCT for a period of time to enable the components of the lamp to stabilise, and to determine when the aging rate has stabilised. If measurements are made of the lamp output over time at a constant

electrical operating current or voltage, the data may be used to estimate the change in the lamp calibration with later use.

#### 5.4.5 Recalibration

Every lamp is an individual. Although an estimated aging rate may be determined as indicated above, the lamp may change suddenly for no apparent reason. The best method to monitor the behaviour of a single lamp is to keep records of every use and to check for any change in the electrical behaviour of the lamp. In a laboratory where accuracy is paramount, measurements should be made with more than one standard to ensure agreement between measurements. When the difference between standards becomes larger than acceptable, consideration must be given to a recalibration of the standards. If the lamps have been used extensively and are changing rapidly, it may be necessary to discard the lamp. However, since standard lamps are costly, sometimes difficult to obtain, and the calibration process is costly, a recalibration of existing lamps should be a first consideration.

If the standard lamps are extensively used in the measurement laboratory, consideration should be given to using the purchased calibrated sources as reference measurement standards to calibrate a set of working measurement standards that can be calibrated in the measurement laboratory and used for internal calibrations.

### 6. Spectroradiometry

One of the basic characteristics of a radiant source is the spectral distribution of the radiant output. As we have seen, even the accurate determination of photometric and colorimetric quantities requires knowledge of the spectral distribution of the radiation. The spectral distribution of the radiation in each of the geometrical quantities that we have discussed may be determined. Since the spectral techniques will be similar in each case, we will illustrate the methods using spectral irradiance as the example. In addition, since spectral irradiance can be used for a large majority of the calibration requirements, it is the most commonly available facility in a laboratory. Several references available for more information concerning spectroradiometric measurements are (CIE 063-1984, Grum & Becherer, 1979, Kostkowski, 1997).

The basic instrument for spectral measurements is a spectroradiometer, which is an instrument for measuring radiometric quantities in narrow wavelength intervals over a given spectral region (CIE S017, 2011). An example of a spectroradiometer measurement configuration for spectral irradiance measurements is shown in Figure 15. The monochromator shown is a grating instrument with a prism predispersor input to remove the high-order wavelengths that would be passed by the grating components. There are two detector output ports with manual switching that allows several wavelength regions to be measured during one measurement set. Detectors such as photomultiplier tubes (PMT), and Si, Ge, InGaAs and InSb photodiodes are available to allow measurements covering the wavelength range from 200 nm to 2500 nm. Various manually interchangeable gratings with different dispersions and blaze are available for use in measuring the same wavelength range. The purpose of the conical radiation trap is to remove radiation emitted in the back direction from scattering back into the input. The aperture defines the spatial region at the spectral irradiance source from which radiation is accepted at the input to the integrating sphere.

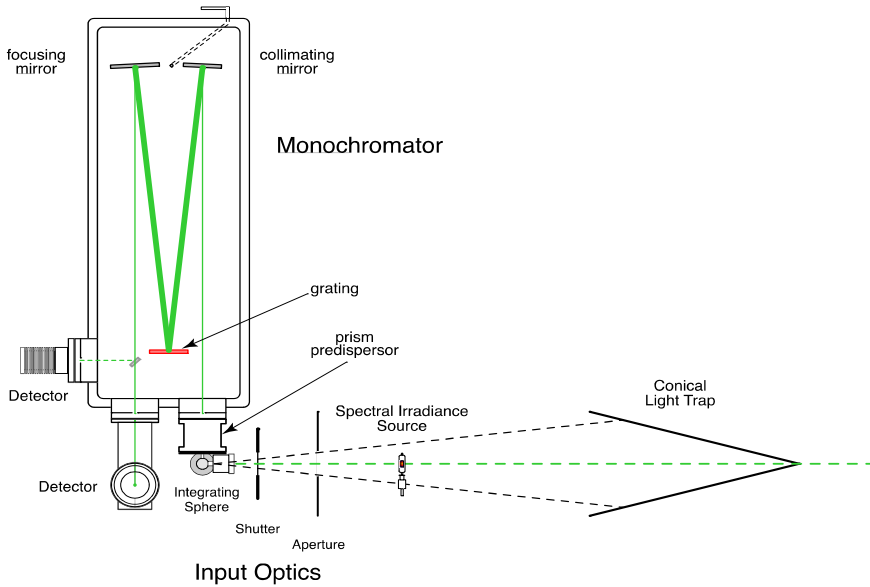


Fig. 15. Spectroradiometer and spectral irradiance measurement

A spectroradiometric measurement system may be considered to be composed of seven basic components (Figure 16): i) the radiation source, ii) the input optics that couples the desired radiation into the monochromator, iii) the monochromator, iv) the output optics that couples the radiation from the monochromator into the detectors, v) the detector, vi) a measurement control system, and vii) some signal processing equipment.

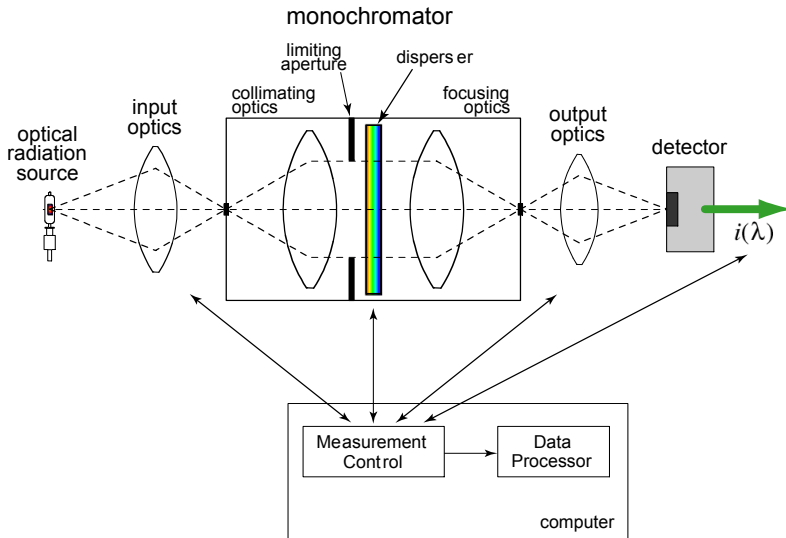


Fig. 16. General configuration for spectroradiometric measurements

For simplicity, in Figure 16 and other figures, lenses (refractive elements) are used to show the focusing of radiation. In most cases, mirrors (reflective elements) are preferable since they will focus all wavelengths of radiation to the same physical spot, whereas lenses do not. For example, the monochromator in Figure 15 uses mirrors in the collimating and focusing optics.

### 6.1 Monochromator

The basic components of a monochromator are shown in Figure 16. The disperser may be a grating or a prism. The limiting aperture serves to define the size of the radiation beam that passes through the monochromator, and limits the size and shape of the input beam that passes through the monochromator. The collimating optics transforms the diverging input beam from the input slit into a parallel beam of radiation incident upon the disperser. The focusing optics focuses the dispersed beam from the disperser onto the output plane at the position of the output slits. The optical and wavelength dispersion properties of a monochromator are best understood by noting that there are two kinds of optics in a monochromator. One kind forms images of the source (input slit) at the output plane (output slit), as shown by the dashed lines inside the monochromator of Figure 16. The optical magnification of the system is usually one, so that the image of the input slits produced at the output plane is the same size as the input slits. The second kind of optics disperses the radiation. For gratings and prisms, this dispersion is spatial with the beams of different wavelengths effectively propagating into different directions from the disperser. As a result, there are many images of the input slits produced at the output plane, one for each of the wavelengths present in the input radiation. Due to the finite width of the input slit and the image, and the continuous wavelength range of most radiation sources, the images of different adjacent wavelengths will partially overlap and there will be a range of wavelengths that are passed through the output slit. This effect gives rise to the bandwidth property of a monochromator.

### 6.2 Spectroradiometer input optics

The purpose of the spectroradiometer input optics is to couple the desired radiation flux from the source into the monochromator as efficiently and accurately as possible. The type of coupling used should take into account the characteristics of the complete system that influence the accuracy of our measurements. Some issues to consider are:

- i. Any non-uniformity in the transmission of any of the optical components, or in the spatial responsivity of the detector, will cause an incorrect 'weighting' of different parts of the spatial radiation distribution. This non-uniformity could depend both upon the direction of the incident radiation, as well as the position of the input radiation. To reduce this effect, the radiation that is input to the monochromator is usually made as spatially uniform as possible.
- ii. Some of the system components will be affected by the polarization of the radiation. This is particularly important when using grating monochromators. The responsivities of some detectors are also affected by the polarization of the incident radiation. This effect will cause an incorrect 'weighting' of different polarization components of the radiation beam. These issues can be resolved by measuring the radiation separately for each polarization, or by spatially mixing the input radiation to provide a uniformly polarized beam.

- iii. Imperfections in the optical components will cause scattering of some of the radiation from its ideal trajectory. This causes both a loss in signal and stray radiation problems.
- iv. The efficiency of our optical design and components will affect our signal-to-noise and accuracy. Since the signals we are usually measuring with a spectroradiometer are small, one of the prime considerations is to collect as much radiation as possible, and to use it as efficiently as possible, without undue loss anywhere in the optical path. To this end, we should use as much of the optical system, particularly that of the monochromator, as possible. The basic ideas of filling the input optics of the monochromator are shown in Figure 17.

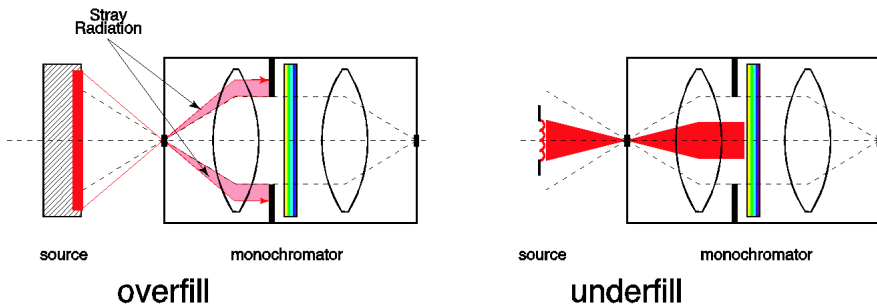


Fig. 17. Source size and monochromator optics

If the radiation from the source overfills the monochromator optics, the overflow radiation will become stray scattered radiation within the monochromator. The problem with this stray radiation is that radiation of wavelengths other than those we are measuring can enter our detection system and cause an erroneous addition to our output signal. A small source, or a source placed at a large distance from the monochromator, will underfill the monochromator optics. In addition to being inefficient, this underfilling of the optics is susceptible to any non-uniformity in the optics. Any motion of the source will cause the radiation to pass through a different part of the spectroradiometer optics, which may result in a different response from the monochromator. This will become more of a problem if another source is used to calibrate the spectroradiometer. This other source will probably be a different size and therefore pass through different optical paths in the system.

- v. An input optic lens or mirror system that collects more radiation from the source and is sized to fit the monochromator input is often used. The schematic of Figure 16 shows this configuration. In this case the filament of the lamp is focused onto the input slits of the monochromator. There are two problems with this arrangement: 1.) Since an image of the filament is produced on the slits, any vibration of the radiation source or spectroradiometer components will cause a dramatic variation in the amount of radiation that passes through the slits and enters the monochromator. 2.) The spatial structure of the radiation at the input slits will be imaged onto the output slits of the monochromator, and then into the detector. If there is any non-uniformity in the detector spatial responsivity, our output signal will not be representative of the amount of input radiation.

There are, however, instances when we do use this type of input optics. These are usually when the source is a uniform source, such as a Lambertian diffuser, or when we

- are imaging the radiation emitted from a certain portion of a ribbon filament of an incandescent lamp. In both of these instances, the source has very little spatial structure.
- vi. In many radiometric measurements, the input system involves a diffuser, either a flat plate (Figure 17) or an integrating sphere. While this may cause a considerable loss of radiation, it does reduce the errors caused by geometrical structure in the radiation source, component vibration, polarization effects, and any spatial non-uniformity in the monochromator optics and/or detector responsivity. The input optics to the monochromator shown in Figure 15 is an integrating sphere, together with a baffle assembly. More detail for this configuration is shown in Figure 18.

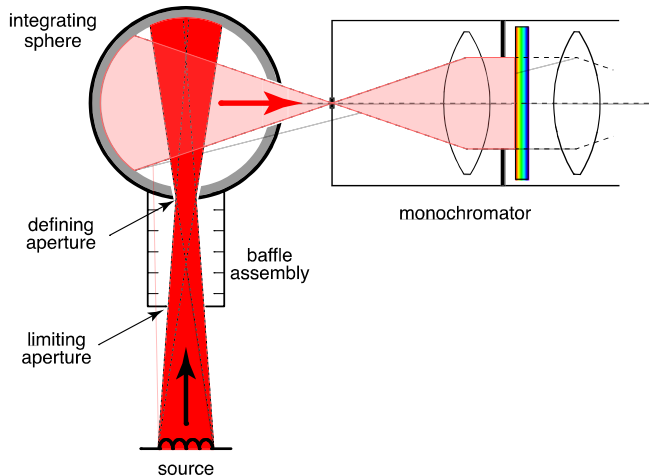


Fig. 18. Integrating sphere input to monochromator

The input port to the sphere is the area that defines what flux is to be measured, so it acts as part of the defining optics for the radiation source measurement. The required distance is that between the source and the defining aperture at the sphere input. The size of the sphere and its ports should be arranged such that the radiation that enters the monochromator has had more than one 'bounce' within the sphere. Figure 18 shows a configuration such that the region of the sphere wall that emits radiation into the monochromator does not overlap with the region of the sphere wall that is directly irradiated by the source. Integrating spheres accept radiation from all directions and emit radiation into all directions at any open port of the sphere. This must be controlled to avoid stray radiation from entering the sphere and to prevent an overflow of the input optics to the monochromator as was shown in Figure 17. To prevent an overflow at the monochromator, the size of the sphere output port and the distance of the sphere output port from the monochromator input slits can be adjusted such that the optics of the input to the monochromator are just filled, as shown in Figure 18.

A baffle assembly may be used at the input to the integrating sphere to reduce unwanted stray radiation from entering the sphere. Apertures and baffles must be considered carefully, since the edges of the aperture or baffle, no matter how thin, will reflect significant radiation, especially from the direct input beam, into the measured portion of the input. The baffle assembly shown in Figure 18 includes only one baffle, the limiting front aperture, that



directly views the input radiation from the source. The baffles interior to the assembly are only used to prevent scattered radiation from the interior walls of the baffle assembly from re-entering the input beam to the sphere. The limiting aperture is large enough that it does not define, or vignette, the final size of the input radiation beam that enters the sphere. Its purpose is to limit as much as possible the amount of off-angle stray radiation that enters the integrating sphere.

The multiple reflections at the interior of the sphere cause a change in the relative spectral distribution of the sphere by a factor of  $\rho(\lambda)/(1-\rho(\lambda))$ , where  $\rho(\lambda)$  is the reflectance of the material of the sphere wall (CIE 084, 1989). Any irregularities or changes in  $\rho(\lambda)$  with wavelength will become amplified in the sphere output radiation. With the typical white diffusers used in integrating spheres, such as PTFE or BaSO<sub>4</sub>, this is usually observed in the lower wavelengths below approximately 400 nm. If a more uniform behaviour with wavelength is desired, a compromise is to reduce the reflectance  $\rho(\lambda)$  from its high values near 99% to approximately 85% by addition of dark absorbers into the coating.

In summary, although the integrating sphere input configuration has several disadvantages, the advantages that it provides in reducing the spatial non-uniformities of sources and detectors outweigh the disadvantages. The disadvantages may be mitigated by careful consideration of each of the components of the system when assembling a measurement system.

### 6.3 Spectroradiometer output optics and detector

The purpose of the output optics is to couple the dispersed radiation from the monochromator, at the monochromator output slit, onto the detector. To obtain the maximum signal-to-noise at the detector, the detector is often placed as close to the output slit as possible in order to collect all the radiation. If the detector must be placed at some distance from the output slits of the monochromator, some optics will be needed to collect the radiation and re-image it onto the detector. Note that if the beam is focused too intensely onto the detector, local overloading of the conversion process from photons to the output (electrical) signal may occur, causing non-linearity in the detector response.

The spectroradiometer shown in Figure 15 is designed to enable various detectors to be used depending on the wavelength range of interest. Each of these detectors requires different output optics of the spectroradiometer to couple the radiation from the monochromator to the sensitive area of the detector. For the VIS and NIR wavelengths, a silicon (Si) photodiode and a InGaAs photodiode can be mounted directly behind the slits. For the UV and VIS wavelengths, a photomultiplier (PMT) is used with a lens assembly to focus the radiation from the slits onto the PMT cathode. The assembly for the infrared (IR) detectors (Ge and InSb) is a bit more complex since these detectors are often cooled with liquid nitrogen and prefer to be operated in an upright position. In addition, their sensitive areas are usually quite small. An elliptical mirror may be used to change the direction of the radiation beam, to focus the radiation onto the detector, and to avoid the problem of a change in the focal position with wavelength (Gaertner & Boivin, 1995).

### 6.4 Measurement procedures

The basic concepts of a measurement and the associated calibration procedures are the same for spectroradiometric measurements as presented in Section 5 above. A measurement is a

comparison between our unknown spectral quantity and a similar quantity of known magnitude. The comparison device or transfer device is now a spectroradiometer, which is a little more complex than the single detectors or photometers we discussed. The input optics to the spectroradiometer must define the geometric quantity that we wish to measure, but the spectroradiometer itself only puts out a signal, such as the current from the detector, which is related to the radiant flux at the input to the spectroradiometer. The basic measurement Equation (34) now becomes a function of wavelength, to be applied at each wavelength that is measured.

$$S^T(\lambda) = \frac{i^T(\lambda)}{i^S(\lambda)} \cdot S^S(\lambda) \cdot \frac{T^S(\lambda)}{T^T(\lambda)} \quad (39)$$

The term  $T(\lambda)$  indicates the effect of all the components of the spectroradiometer, from the input optics to the detector (Figure 16), upon the input source flux  $S(\lambda)$ . Usually we assume that the effect of  $T(\lambda)$  upon the measurements is the same for both the test and the standard sources, which reduces Equation (39) to the familiar form of Equation (34) for each wavelength. In spectroradiometric measurements the terms that must remain constant between the measurement of the standard and the test quantity are quite complex. Characteristics of the monochromator such as the large peaks in the grating reflectance and the reproducibility of the wavelength scale place greater demands on the ability of the instrument to reproduce its settings. The effects of these two particular characteristics can be mitigated by performing the measurements on both sources at each wavelength, before adjusting the monochromator to the next required wavelength. This requires that switching the monochromator between the test source and the standard source can be done in an accurate and reproducible manner. In addition to requiring the reproducibility of the monochromator, the entire optical system must 'evaluate' the flux in the same manner for the two measurements. This means that nothing in the behaviour of  $T(\lambda)$  should depend on the magnitude of  $S(\lambda)$ , nor its initial direction or position within the defined path.

## 7. Primary radiation sources and calibration chains

In Section 3 we introduced the concept of a metrological traceability chain that enables our measurements to be accurate and reproducible worldwide. The primary measurement standards used for optical radiation measurements may be traceable to either detector standards or source standards that provide a procedure by which the reference quantity is the definition of the measurement unit through its practical measurement. The two sources that are typically used in the source-based method are (1) blackbody radiators, whose output is calculated from the Planck equation when the temperature of the blackbody is known (Mielenz et al, 1990), and (2) synchrotron radiators, whose output can be determined from the calculable radiance of accelerating charged particles (Ulm, 2003). The detector-based method depends upon the measurement of absolute quantities of radiant flux by absolute radiometers (Boivin & Gibb, 1995), which compare the heating caused by the absorption of the radiant flux with the heating caused by a known amount of electrical power. A comparison of the basic optical radiation calibration chain (Gaertner, 2009) for each of these methods, using a high temperature blackbody radiator source (HTBB), is given in Figure 19.

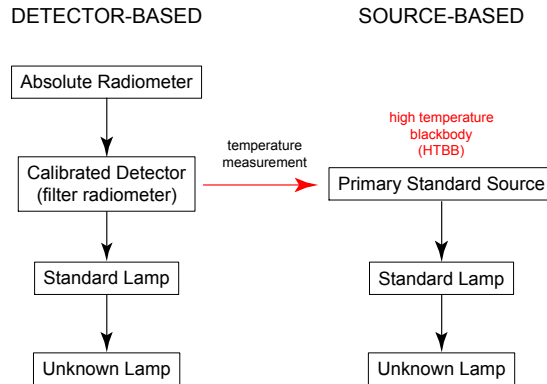


Fig. 19. Detector-based and HTBB Source-based calibration chains for optical radiation measurements.

This calibration chain indicates that HTBB primary source based calibrations are based upon primary detector measurements. The absolute temperature of a HTBB must be determined by the absolute measurement of the radiance of the HTBB. This is done using filter radiometers that measure the radiance or irradiance of the HTBB within a limited wavelength range. The filter radiometers used for irradiance measurements (Boivin et al, 2010) are basically constructed from detectors with a wavelength-limiting filter that allows the spectral responsivity of the filter radiometer to be accurately calibrated from measurements traceable to the absolute cryogenic radiometer. The measurement of the radiance of the HTBB is performed using the geometry described in Section 4.1.7.4, with a calibrated filter radiometer as the receiver. After the temperature of the HTBB is determined, the spectral radiance of the HTBB may be calculated for the desired wavelength range using the Planck equation. The HTBB may then be used for spectral irradiance calibrations using the geometry described in Section 4.1.7.4. With HTBB operating temperatures from 2500 K to 3500 K, known spectral irradiances may be produced from wavelengths of 200 nm to over 2500 nm to calibrate standard incandescent lamps, such as the 1000 W FEL lamps that are commonly used as spectral irradiance measurement standards. Since the radiation output of an incandescent lamp is quite low in the UV, deuterium arc lamps are often used in the wavelength range from 200 nm to 350 nm. These can be calibrated using HTBB sources operating in the upper range (3400 K to 3500 K) of their temperature limits.

Whereas the transfer from detector-based to source-based measurements is very useful for spectral radiation measurements, the calibration of sources that are used for photometric measurements is best performed using filter radiometers directly to calibrate the standard incandescent lamps (Gaertner et al, 2008). For this purpose, the filter radiometer is constructed with a combination of filters that results in a filter radiometer with a spectral responsivity that approaches the photometric functions as closely as possible.

## 8. Uncertainties

A measurement is not complete until an estimate of the uncertainties has been made and an uncertainty budget prepared. The GUM (JCGM, 2008a) and related documents (JCGM, 2009)

are excellent references. In addition, the CIE has published an extensive document on uncertainty determinations (CIE 198:2011).

The measurement configurations discussed in the preceding sections have served to define the typical types of optical radiation measurements. This means that we will be measuring radiation using internationally defined quantities and units. This requires that the actual measurement configurations and equipment that we use must adhere to the internationally accepted definitions.

There are many factors that will influence the results of our measurements. In addition to the difficulty of configuring our apparatus to measure the quantities we wish to measure, there will be influences of time, temperature, humidity and many other often unknown or unexpected factors upon the equipment we use to give us our results. The consequences are that the results we obtain are not exactly what we wish to obtain, nor are they what we would like to claim them to be. Our only recourse is to try to understand the behaviour of our equipment and how many of the conditions in our laboratory can influence the results. In many cases this will mean a deliberate attempt to change these variables to determine the subsequent change in our results. By this means we will come to know the influences that change the results by significant amounts, and are therefore important influences to control. If we cannot control them, we will at least have an idea as to how to correct our results to account for the error caused by them, or at minimum, the amount of an uncertainty to apply to our results.

The evaluation of measurement uncertainties has been divided into two different types: Type A and Type B (JCGM 100:2008, JCGM 104:2009, JCGM 200:2008).

### 8.1 Type A uncertainty evaluation

The Type A evaluation of a component of our measurement uncertainties is done by a statistical analysis of the measurement quantity values that are obtained by repeated measurements under the same defined measurement conditions. Each measurement value under these conditions is different from the previous measured value in a random manner such that the next measured value cannot be predicted exactly from the previous value. These uncertainties are usually analysed using Gaussian probability density functions, such as shown in Figure 20. If many repeated measurements are made, the distribution describing the quantity  $x$  will approach a curve similar to the black curve labeled 'distribution of data'. From these measured values,  $x_{mean}$ , the mean value of  $x$ , and  $\sigma_{data}$ , the standard deviation of the distribution of the measured data, can be calculated. Any single measured value of  $x$  is expected to be found within this distribution.

The accuracy or uncertainty with which we know both  $x_{mean}$  and  $\sigma_{data}$  depends upon how well we know the 'distribution of data' curve, which will depend upon how many data points we have taken. The accuracy of  $x_{mean}$  is described by the standard deviation of the mean,  $\sigma_{mean}$ , which is calculated from the datapoints by:

$$\sigma_{mean} = \sigma_{data} / \sqrt{n} \quad (40)$$

where  $n$  is the number of datapoints. The red curve labeled 'distribution of mean' in Figure 20 shows the uncertainty with which the mean value  $x_{mean}$  is known from  $n$  measurements

of  $x$ , all of which are assumed to be part of the distribution shown by the 'distribution of data' curve, and as calculated for a Gaussian distribution using equation (40). The value used for  $n$  was 10 in Figure 20.

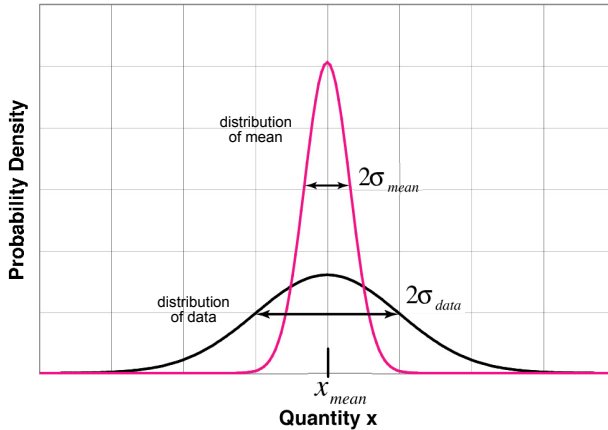


Fig. 20. Gaussian probability density functions for Type A evaluation of uncertainties.

The distinction between  $\sigma_{data}$  and  $\sigma_{mean}$  is important for the estimation of the Type A uncertainties in the calibration and use of optical radiation detectors, such as photometers and spectroradiometers. During the calibration of the detector, a total of  $n$  measurements may be performed, resulting in a mean calibration value of  $x_{mean}$  with a Type A uncertainty of  $\sigma_{mean}$ . When the detector is later used for a single measurement of quantity  $x$ , using the calibrated value obtained from  $x_{mean}$ , the Type A uncertainty in this single measured value will be given by  $\sigma_{data}$ , the standard distribution of the probability distribution that describes the repeatability of the detector. The uncertainty in the calibration of the detector, which includes the  $\sigma_{mean}$  obtained during the calibration, now becomes an uncertainty with a Type B uncertainty evaluation for the single subsequent measurement.

## 8.2 Type B uncertainty evaluation

The Type B uncertainty evaluation is defined as an evaluation of a component of measurement uncertainty determined by means other than a Type A evaluation uncertainty (JCGM, 2008b). This type of evaluation is used with systematic errors, which are themselves associated with the fact that a measured quantity value contains an offset from the true quantity value. These offsets have many origins, such as those indicated in Section 8.0 above. Determining actual and potential systematic errors requires critical evaluation of our measurement configuration.

As we have seen, optical radiation measurements involve the determination of the flux produced by a source into many different geometrical configurations. As a result, these measurements will require the use of working standard sources to either calibrate our detection system or to compare with our test source using the detection system. We noted in Section 4.2.5 the importance of comparing sources as similar as possible to reduce the spectral errors and uncertainties in photometric measurements. In Section 6.4 we discussed

the importance of accounting for the potential differing responses of the spectroradiometer ( $T(\lambda)$ ) to the different sources. As may be expected, this 'like-with-like' principle is also applicable to geometrical properties and in signal size comparisons. Therefore we should use a working standard lamp that is as similar in relative spectral output, geometrical size and shape, and flux output as the test lamp we wish to measure. By using all these techniques we will stack as many odds in our favor as possible.

However, as discussed below, there will still be aspects of our measurement that will cause us some errors and uncertainties.

### 8.2.1 Radiation sources

Some of the properties of optical radiation measurement standard sources were introduced in Section 5.4. The following sources of error and uncertainty must be considered for both the working standard and the test sources:

1. The calibrated value from the Calibration Laboratory will indicate an uncertainty in the value that has been placed on the working standard source. This is your minimum uncertainty, since you cannot do anything about it.
2. The source calibration is given for defined electrical operating conditions. The uncertainty with which you can reproduce the defined current (or voltage) through the lamp will cause an uncertainty in the amount of radiation that you are receiving from the lamp. As indicated in Section 5.4.3 above, this can be a large effect. This uncertainty will be dependent upon the calibration of the voltmeter and the standard resistor that are used to set the lamp current. The ability of the power supply to fine-tune the electrical values to obtain the defined values will affect the lamp output. In addition, the short-term fluctuations in the electrical output of the power supply will add to instability in the source radiation.
3. The calibration value is true for a defined geometric alignment of the source. This usually means the reference point in the lamp filament and the direction from the plane of the lamp filament for which the calibration was made. The source output may be different in different directions from the filament. If this information is known, or if you can measure it, an uncertainty may be calculated due to your uncertainty in exactly reproducing the original calibration conditions.
4. The intensity of the source may vary with distance from the lamp, or the irradiance due to the source may not follow the inverse-square-law. To test this for your measurement configuration, and for accurate results, a series of measurements at different distances from the source should be performed to test the inverse-square-law behaviour of the lamp (see Section 4.1.7.1).
5. The irradiance from the source at the detector may not be spatially uniform. Every measurement is actually an average irradiance over the full area of the acceptance aperture of the detector. Any difference in the size of the spatial region of the source irradiance field between that calibrated by the Calibration Laboratory and that used in your measurements will result in an error due to this spatial non-uniformity.
6. The radiation output of incandescent lamps changes with use. This aging effect can be approximated to enable a correction to be made, but there will be an associated uncertainty with this correction. This aging effect for each lamp may be approximated from measurement data that can be made when the lamp is originally aged before calibration (Section 5.4.4).

### 8.2.2 Radiation detectors

The complexity of photometers and spectroradiometers gives rise to many interesting sources of error.

1. The failure of a photometer to match the desired photometric or colorimetric functions has been discussed in Section 4.2.5. The uncertainty in the spectral mismatch correction factor will need to be determined from the estimated uncertainties in the spectral distribution functions for the contributing sources and the photometer.
2. The response of the detector used in the photometer or spectroradiometer to input flux signal size must be linear to allow comparison of sources with differing amounts of radiation output. This effect can also include the associated electronics.
3. If the geometric distribution of the output of the lamps varies with angle or position, and if measurements of high accuracy are required, any spatial non-uniformity in the responsivity of a photometer will have an effect. The use of integrating spheres to mitigate this source of error was discussed in Section 6.2.
4. If the number of digits available in the output of the voltmeter used to measure the signal size is small, the measured signals will be subject to a digitizing error.
5. Monochromator wavelength errors. There are two possible components to this error: the deviation of the indicated monochromator wavelength from its actual value, and the ability of the monochromator to reproduce this error. The reproducibility error may give rise to a random error that can be determined using a Type A uncertainty evaluation if enough repeatable measurements can be made. The effects of an error in the monochromator wavelength calibration will predominantly depend upon the difference in the relative spectral distribution of the two sources that are being compared. If the sources have the same relative spectral distribution, the wavelength offset will cause the same relative effect for both sources, resulting in no error in calibration.
6. Spectral bandwidth errors. The signal at the spectroradiometer detector is a weighted average over the spectral bandwidth of the monochromator. If the spectral distribution of the source changes rapidly over this bandwidth, much of this spectral information will be lost in the averaging process.

More detail concerning these and other effects may be found in several of the references (CIE 063:1984, Grum & Becherer, 1979, Kostkowski, 1997).

### 9. Conclusion

The measurement of absolute amounts of optical radiation requires careful and detailed consideration of a broad range of physical concepts and practical instrumentation to produce an accurate, reproducible and internationally acceptable result. The basis for this is the internationally defined terminology and definitions of the measurement quantities and accepted units that are useful for optical radiation measurements. We have discussed the five predominant measurement quantities and indicated the measurement configurations necessary to obtain reliable results. The geometrical relationships between these quantities have been discussed with the purpose of allowing measurement standards to be used for several different measurement quantities. Some of the requirements to measure optical radiation in a means that is relevant to the human visual response to radiation have been presented. The characteristics of incandescent sources that are a basis for most measurements have been discussed, together with the sources of errors and uncertainties that these

characteristics present. The basic concepts of the photometers and spectroradiometers that are used to measure the radiation output from these sources have been presented with the aim of providing a reliable foundation for basic metrology and applied measurements. The necessity of international acceptability of goods and services, energy efficiency, and consumer safety have encouraged the development of more accurate, versatile and reliable equipment both to produce a wider variety of radiation sources and better means of determining their suitability to the desired application. The basic concepts and techniques described in this chapter should provide the necessary tools to pursue these goals.

## 10. References

- BIPM (2006), The International System of Units (SI), 8<sup>th</sup> edition. Available from <http://www.bipm.org>.
- Boivin, L.P. & Gibb, K., (1995). Monochromator-based cryogenic radiometry at the NRC, *Metrologia* Vol. 32, No. 6, (1995), pp. 565-570.
- Boivin, L.P., Bamber, C., Gaertner, A.A., Gerson, R.K., Woods, D.J. & Woolliams, E.R. (2010) Wideband filter radiometers for blackbody temperature measurements, *Journal of Modern Optics*, Vol. 57, No. 17, (October 2010), pp. 1648-1660, ISSN 0950-0340.
- CIE 015:2004 Colorimetry, 3<sup>rd</sup> Edition, ISBN 978-3-901906-33-6, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 063 (1984), The spectroradiometric measurement of light sources, ISBN 978-963-7251-23-8, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 069 (1987), Methods of Characterising Illuminance Meters and Luminance Meters, ISBN-978-3-900734-04-6, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 084 (1989), The Measurement of Luminous Flux, ISBN-978-3-900734-21-3, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 127 (2007), Measurement of LEDS, ISBN 978-3-901-906-58-9, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 149 (2002), The Use of Tungsten Filament Lamps as Secondary Standard Sources, ISBN 3-901-906-18-5, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 179 (2007), Methods for Characterising Tristimulus Colorimeters for Measuring the Colour of Light, ISBN 978 3 901906 60 2, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 191:2010 Recommended System for Mesopic Photometry Based on Visual Performance, ISBN 978-3-901906-88-6, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE 198:2011 Determination of Measurement Uncertainties in Photometry, ISBN 978-3-902842-00-8, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE Standard S 010/E:2004, Joint ISO/CIE Standard ISO 23539:2005(E) / CIE S010/E:2004 Photometry - The CIE System of Physical Photometry, Commission Internationale de L'Eclairage, Vienna, Austria.
- CIE Standard S 014-2/E:2006, Joint ISO/CIE Standard ISO 11664-2:2007(E)/CIE S 014-2/E:2006 CIE Colorimetry-Part 2: Standard Illuminants for Colorimetry, Commission Internationale de L'Eclairage, Vienna, Austria
- CIE Standard S 017/E:2011, ILV: International Lighting Vocabulary, Commission Internationale de L'Eclairage, Vienna, Austria.



- Gaertner, A.A. & Boivin, L.P., (1995). Some problems in realizing an infrared spectral-irradiance scale from 1500 nm to 2400 nm at the NRC, *Metrologia* Vol. 32, No. 6, (1995), pp. 615-619.
- Gaertner, A.A., Bamber, C., Boivin, L.P. and Chrysler, M. (2008) NRC photometer design for the realisation of a luminous intensity scale, CIE publication x033:2008 Proceedings of the CIE Expert Symposium on "Advances in Photometry and Colorimetry", 7-8 July 2008, Turin, Italy.
- Gaertner, A.A., (2009). Spectral irradiance calibrations at the National Research Council of Canada (NRC), *Journal of Modern Optics* Vol. 56, No. 13, (July 2009), pp. 1488-1496, ISSN 0950-0340.
- Gardner, J.L. (2000). Correlated Colour Temperature - Uncertainty and Estimation, *Metrologia* Vol. 37, No. 5 (2000), pp. 381-384.
- Grum, F. and Becherer, R.J., (1979) *Optical Radiation Measurements Volume 1 Radiometry*, Academic Press ISBN 0-12-304901-6 (v.1), New York, USA.
- Höpe, A. & Hauer, K-O. (2010) Three-dimensional appearance characterization of diffuse standard reflection materials, *Metrologia* Vol. 47, No. 3, (June 2010), pp. 295-304.
- JCGM 100:2008, Joint Committee for Guides in Metrology (September 2008), Evaluation of Measurement Data – Guide to the expression of uncertainty in measurement (GUM). Available from <http://www.bipm.org>.
- JCGM 104:2009, Joint Committee for Guides in Metrology (July 2009), Evaluation of Measurement Data – An introduction to the "Guide to the expression of uncertainty in measurement" and related documents. Available from <http://www.bipm.org>.
- JCGM 200:2008, Joint Committee for Guides in Metrology (2008), International vocabulary of metrology – Basic and general concepts and associated terms (VIM). Available from <http://www.bipm.org>.
- Kostkowski, H.J. (1997) *Reliable Spectroradiometry*, Spectroradiometry Consulting, ISBN-0-9657713-0-X, La Plata, Maryland, USA.
- Mielenz, K.D., Saunders, R.D., Parr, A.C. & Hsia, J.J. (1990), The 1990 NIST Scales of Thermal Radiometry, *J. Res. Natl. Inst. Stand. Technol.* Vol 95, No. 6 (December 1990), pp. 621-629.
- Ohno, Y. (1997) Measurement Procedures (Chapter 5) and Photometric Standards (Chapter 3), In *Handbook of Applied Photometry*, Casimir DeCusatis, (133-177, 55-99) American Institute of Physics Press, ISBN 1-56396-416-3, Woodbury, NY, USA.
- Ohta, N. & Robertson, A.R. (2005) *Colorimetry Fundamentals and Applications*, Wiley, ISBN-13 978-0-470-09472-3, Chichester, England
- Ulm, G., (2003). Radiometry with synchrotron radiation, *Metrologia* Vol. 40, No. 1, (February 2003), pp. S101-S106.
- Walsh, John W.T. (1958), *Photometry*, 3<sup>rd</sup> Edition, Constable, London, UK.

# Optical Measurements: Polarization and Coherence of Light Fields

O. V. Angelsky<sup>1</sup>, P. V. Polyanskii<sup>1</sup>, I. I. Mokhun<sup>1</sup>,  
C. Yu. Zenkova<sup>1</sup>, H. V. Bogatyryova<sup>2</sup>, Ch. V. Felde<sup>1</sup>,  
V. T. Bachinskiy<sup>3</sup>, T. M. Boichuk<sup>3</sup> and A. G. Ushenko<sup>1</sup>

<sup>1</sup>*Chernivtsi National University, Chernivtsi,*

<sup>2</sup>*National Technical University of Ukraine "Kiev Polytechnic Institute", Kiev,*

<sup>3</sup>*Bukovonian State Medical University, Chernivtsi,  
Ukraine*

## 1. Introduction

The Chapter is devoted to consideration of metrological aspects of intrinsically interconnected characteristics of light fields, such as intensity, polarization and coherence. Conceptually, all these quantities are derived from the Wolf's coherency matrix [1]. However, new insight on interconnection of them is provided by the novel singular-optical approach [2, 3] predicting existence of important regularities in electromagnetic fields which were early considered as quite random ones. So, phase singularities of scalar (homogeneously polarized), polarization singularities of vector (inhomogeneously polarized) fields, as well as singularities of correlation functions of partially coherent, partially polarized fields constitute specific skeletons, i.e. "bearing" elements of a field. Knowing the loci and characteristics of such elements, one can judge on behavior of a field at its other areas, at least in qualitative manner, but quite reliably [4]. This circumstance opens quite new possibilities for metrology of optical fields and leads to prospective practical applications of new metrological techniques.

We discuss here interconnections of polarization and coherence characteristics of light fields in various manifestations of them, both as one-point and two-point functions. For that, in development of earlier approaches, we show the framework for generalization of polarization metrology on a wide class of combined optical beams assembled as mutually incoherent (or partially mutually coherent) components, which can be orthogonal in polarization. Such generalization provides taking into account partial polarization and associated vector singularities, which can be used in non-destructive optical diagnostics as well as in optical telecommunications with polarization coding.

Further, we represent point-by-point Stokes-polarimetric technique and singular-optical concept of polarization diagnostics. Feasibilities of combined application of conventional interferometry and local Stokes-polarimetry are substantiated.

Important part of this review is devoted to description of feasibilities for experimental measuring of coherence by measuring of spatial polarization distributions of

inhomogeneously polarized fields. We represent the newest metrological tool connected with novel concept of optical currents (optical flows). Namely, we show that some intimate characteristics of complex optical fields with arbitrary degree of spatial coherence and arbitrary degree of polarization may be “deciphered” indirectly, by observation of the influence of such fields on embedded micro- and nanoparticles. This original metrological approach seems to be prospective for development of so-called optical traps and tweezers for manipulation of isolated particles of micro- and nanoscales.

Separate section of the Chapter is devoted to application of local Stokes-polarimetry in diagnostics of biological tissues, in the context of early (pre-clinical) diagnostics of some widespread diseases. We represent both experimental and data processing techniques leading to high-sensitive and reliable diagnostics.

All considered metrological approaches and techniques are original, generated by recently by the members of our team.

Some prospects of further investigations in the direction represented in this Chapter, as well as necessity and possible ways for overcoming some present shortcomings of optical metrology in the field of coherence and polarization, are outlined in the last section.

## 2. Historical outlook

The notion of coherence is the most fundamental concept of modern optics. As it has been shown by E. Wolf [5], this notion is intrinsically connected with other characteristics of light, such as intensity and polarization. One can distinguish between these characteristics mainly in didactic purposes. But in every practically important case we meet the problem of tight, inseparable interconnection of them. So, one cannot define coherence, in part aspiring to associate it with visibility of interference pattern, ignoring for that the states of polarization of superposed beams. Note, that attempts to explain the Young’s interference experiment for “completely unpolarized” light lead sometimes to questionable conclusions [6]. At the same time, the most fundamental definition of polarized light is given just through the measure of mutual coherence of the orthogonally polarized components of a beam. At last, all three mentioned characteristics of a light beam are comprehensively expressed through known combinations of the Wolf’s coherency matrix elements [1].

Incidentally, urge towards to associate coherence just with obvious interference (intensity modulation) effect does not always lead to true understanding the coherence phenomena. It is not enough that interference fringes are absent in superposition of completely mutually coherent but orthogonally polarized beams (it is well-known from the Fresnel-Arago laws and experiments). There are quite new concepts showing the absence of interference effect for superposing two waves of equal frequencies with strictly (*deterministically*) connected complex disturbances even with the same state of polarization. Refined example of this kind was given by L. Mandel in his concept of “anticoherence” [7]. In very simplified form, the Mandel’s concept can be formulated in terms of conventional (static) holography.

As it is well known, a simple thin hologram reconstructs in plus-minus first diffraction orders two conjugate fields producing so-called main and conjugate images [8]. Having reliable techniques for optical phase conjugation, one can try to add two complex conjugated copies of the signal at one plane, to say at a plane of hologram. What is the result? Intensity of superposition of two such waves is determined as

$$I = \langle |a \exp[i(\omega t - \mathbf{k}\mathbf{r} + \varphi_1)] + b \exp[-i(\omega t - \mathbf{k}\mathbf{r} + \varphi_2)]|^2 \rangle, \quad (1)$$

where  $a$  and  $b$  are amplitudes of two conjugated waves that are believed constants (stationary optical fields [9]),  $\omega = 2\pi/T$  is angular frequency of oscillations ( $T$  being a time period) that is the same for both superposed counterparts,  $\mathbf{k}$  is the wave vector,  $\mathbf{r}$  is the position vector of the observation point,  $\varphi_{1,2}$  are the initial phases of two waves, and  $\langle \dots \rangle$  denotes time averaging. It is easy to see that, in contrast with common interference, the temporal multiplier does not vanish in this expression:

$$I = a^2 + b^2 + 2ab \langle \cos[2(\omega t - \mathbf{k}\mathbf{r}) + (\varphi_1 - \varphi_2)] \rangle \equiv a^2 + b^2, \quad (2)$$

so that the averaging results in vanishing the “interference” term. This conclusion is illustrated in Fig. 1, from which one can see that summation of two complex conjugated beams,  $u$  and  $u^*$ , everywhere along the Arrow of time gives rise to the real value of constant magnitude. Thus, interference between deterministically, unambiguously connected two waves of equal frequency and identical state of polarization is absent. Of course, if one implements phase conjugation of one of two waves figuring in Eq. (1), interference effect will be provided due to compensation of the temporal multiplier and, as a consequence, inefficiency of time averaging.

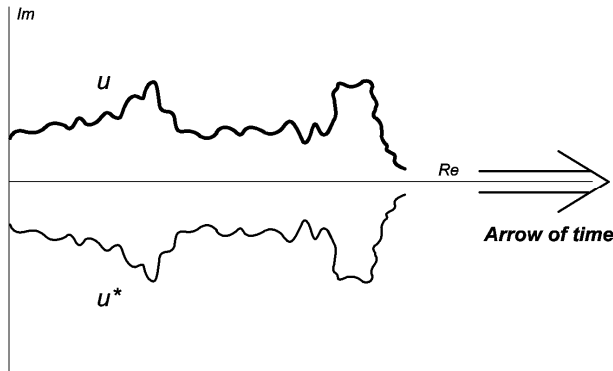


Fig. 1. Superposition of two complex conjugated replicas of a signal results in a real signal of constant amplitude along the Arrow of time, without interference effect.

What is less discussed within the framework of classical (wave) theory of partial coherence is the influence of an “observer” on coherent properties of studied beams. This problem is quite typical for quantum optics [10 - 12] (as well as for quantum physics in general [13]). But the results obtained along last quarter of a century compel to take into account influence of the conditions of observation and detector characteristics on evaluation of the coherence of light in classical approach also. Let us give some arguments for this point.

When we consider interference of two waves with close but non-equal frequencies ( $\omega_1 \neq \omega_2$ ,  $\Delta\omega \ll \bar{\omega}$ ), we observe moving intensity fringes. This effect is widely used in the optical

heterodyning (optical nonlinear mixing) technique. If the “register” is stationary and possesses low temporal resolution, averaging over large enough temporal interval results in smoothing of interference fringes, so that visibility  $V \rightarrow 0$ ; we conclude that two waves are mutually incoherent. But, on the other hand, using a “register” with higher temporal resolution, which moves in the direction and with speed of propagation of moving intensity waves, we recognize the same two waves as mutually coherent that *form* observable interference fringes in moving coordinates! To say, one can register a hologram from such waves at the properly moving “register” (though this task is not simple in practice).

Once more example related to the problem of interest is the influence of *spectral* resolution (or, more strictly, inhomogeneous spectral sensitivity) of the detector on our conclusions concerning *spatial* coherence of the elaborated optical field. If one uses in the classical Young’s two-pinhole interference arrangement a source with uniform spectral density, then a detector with uniform spectral sensitivity “evaluates” the field as completely spatially incoherent for arbitrary sampling points at the beam cross-section due to superposing numerous scaled in wavelength spectral interference patterns, so that minima of intensity (of the field of homogeneous spectral distribution!) are absent in the resulting pattern. *Such detector is “blind” to spatial coherence of such optical field.* Nevertheless, the human eye detects spatial coherence of a field due to non-uniform spectral sensitivity of visual receptors and inhomogeneous spatial distribution of *colors* over analyzed field. Of course, the same is true for evaluation of *temporal* coherence, to say, in the arrangement of the Newton’s rings in white light.

Less trivial case has been considered by Wolf [14, 15, see also numerous useful references herein] in the context of induced spectral changes resulting in remarkable transformations of temporal coherence of polychromatic optical fields due to the presence of material *intermediary*, as diffraction or scattering object.

The next, and more closer to our consideration, example – *pseudodepolarization* [16] (in modern terminology, “global” depolarization [6]) resulting from stationary scattering of laser radiation in multiply scattering media, such as turbid media, multi-mode waveguides, the most of natural objects, including biological ones. Here the role of detector becomes fundamental. Really, the universal approach to determine all polarization characteristics of a field (both the state of polarization and the degree of polarization [17]) consists in Stokes-polarimetry of the analyzed field. For that, Stokes-polarimetric analysis gives quite different results for local and “global” (space-averaged) measurements. So, the point-wise measuring Stokes parameters shows complete (unity) degree of polarization, but the state of polarization changes from point to point. Space averaging over ten and more speckles shows seeming depolarization. This case is the central subject of study of vector singular optics [18].

The mentioned examples illustrate, in part, importance of taking into account of temporal and spatial resolution of the used detector, as well as a choice of the coordinates (stationary or moving) for metrological estimation of coherence and polarization of light, and even understanding (*definition*) of these phenomena.

Let us briefly highlight another two positions that are important for our consideration. Firstly, as it has been pointed out by I. Freund [19], we do not ready to investigate experimentally the problem of coherence and polarization of optical light in general, 3D

case, when paraxial approximation violates and one cannot neglect any of three Cartesian coordinates for description of behaviour of the electric vector. As an example, Freund references the study of polarization of relict - cosmic microwave radiation (CMR) [20] that is believed to be almost *isotropic* (nondirectional) [21]. Measuring the Stokes parameters for such radiation is, to all appearance, not well-grounded. It is the same that measuring the Stokes parameters in various directions from interior of a cloud of light-scattering small particles, what approach is not generally accepted [22, 23]. Nevertheless, it is the only what we have!

The second example concerns directly to one of the problem discussed in the following sections of this overview, *viz.* so-called *optical currents (flows)* [24]. Though it is prematurely now to solve comprehensively this problem, especially in experimental aspect, it is clear that micro- or nanoparticles serving for diagnostics of inhomogeneously polarized and partially coherent optical field [25 - 30] affect this field as absorbing and retransmitting particles with their own characteristics, so that the state of a field, in general, changes under influence of such secondary radiators.

Pronouncing call of the times in the topic under consideration consists in involving the ideas, approaches and techniques of Singular Optics [2]. It is seen, in part, from recent important review [3] devoted to the structure of partially coherent optical fields. Investigations in this scientific domain have been actually stimulated, in part, by papers [31 - 34]. For that, many usual and new results of the theory of partial coherence and partial polarization become essentially urgent just in the singular optics concept. On the other hand, there are good reasons to wait that attracting the fundamentals of the theory of partial coherence will lead to development of new practical applications of singular optics. So, two mentioned areas are of "mutual benefits".

Moreover, as it has been argued in papers [30, 35, 36], "*Usual beam parameters either characterize a beam 'in a whole' (power, momentum, beam size and divergence angle) or describe its 'shape' via certain spatial distributions (amplitude, phase, polarization state, etc.)... Usual beam parameters provide only rough and, sometimes, distorted picture of internal processes that constitute a real 'inner life' of a light beam. These processes are related to the fundamental dynamical and geometrical aspects of light fields, and are associated with the permanent energy redistribution inside the beam 'body', which underlies the beam evolution and transformations. The internal energy flows provide a natural and efficient way for 'peering' into the light fields and studying their most intimate and deep features.*" It is of interest, in the context of this review, to correlate this statement with the Wolf's methodology of observable quantities that is the most influential concept of physical optics since 1954: "optics of observable quantities, such as correlation functions and averaged in time intensities" [17]. Paradoxical contradiction between two undoubtedly true statements is apparent. Really, this contradiction is just eliminated as one takes into account that internal energy flows ("optical currents" [24]) may be revealed only by carrying out the experiments with observable quantities. Similarly, vibrational phase [18] of an optical wave is unobservable, while its difference with a phase of coherent reference phase is liable to registration just as the phase of the mutual coherence function of two waves. More globally, two mentioned approaches are complementary, being intimately interconnected, similarly to approaches of statistical physics and thermodynamics.

To facilitate understanding the material of the following sections, let us briefly remind the main light beam characteristics involved in our consideration. The key notion of *geometrical optics* is a ray; a bundle of rays constitutes an optical beam. The only information provided by rays is the direction of propagation of light. More diverse, more fundamental and more informative parameters of light are derived within the framework of wave optics. There are wave amplitude and phase (both quantities are unobservable directly though lying in the basis of theory [17]); polarization that is determined by the amplitude ratio and phase difference of orthogonal components of a beam; the degree of polarization is determined by correlations between these components; coherence parameters characterizing concordance of light disturbances at two (or more) spatial-temporal points of light field. These base characteristics defined in seminal book [17] as well as in numerous other handbooks in optics, are added by the Poynting vector determining the direction of propagation of light beams in terms of electrodynamics, and by more complex polarization parameters than simple azimuth of polarization and ellipticity inherent in elementary wave optics. In our study, much attention will be paid to the Stokes parameters, polarization parameters intrinsically connected with both energy and coherent characteristics of light beams. These notions will be specified in the corresponding sections of the chapter.

Organization of the review is not quite usual. The most of review papers written up to now on the topic of interest are of theoretical nature, sometimes with valued simulation background. Relatively less attention is paid to experimental aspects of the problem. Partially, this gap is filled by recent books and book chapters written by the authors of this review [37, 38]. But the scope of experimental results rapidly grows, and this outlook does not repeat our previous papers. So, we represent here several independent experimental current views on the problem of metrology of coherent and polarization properties of optical fields, with especial accent on singular optical prerequisites and consequences of our experimental approaches. In spite of forced incompleteness of this consideration, the authors hope that it would stimulate further experimental investigations in this field and lead to broadening of practically significant applications of *Singular Optics of Partially Coherent and Inhomogeneously Polarized Optical Fields*.

### **3. Investigation of optical currents in completely coherent and partially coherent vector fields**

In this section we present the computer simulation results on the spatial distribution of the Poynting vector and illustrate motion of micro and nanoparticles in spatially inhomogeneously polarized fields.

The Poynting vector  $\mathbf{S}$ , is defined, in its simplest form, as the vector product of the vectors of electric field,  $\mathbf{E}$ , and magnetic field, *viz.*  $\mathbf{H}$ ,  $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ . This form is referred to as the Abraham form of the Poynting vector. By definition, the Poynting vector *for a plane wave* is perpendicular to vectors  $\mathbf{E}$  and  $\mathbf{H}$ , being representing the energy flux (in  $\text{W}/\text{m}^2$ ) of an electromagnetic field. As it will be seen from the following consideration, in light fields with complex wave fronts and with variable in space (inhomogeneous) state of polarization the Poynting vector can exhibit much more sophisticated behavior, being also changing from point to point at the beam cross-section and leading to new applications of light.

The use of small particles for diagnostics of microstructure of light is widely used approach [25 - 30], but mainly in approximation of complete coherence of an optical field. Here, the influence of phase relations and the degree of mutual coherence of superposing waves in two-wave and four-wave configurations on the characteristics of the microparticle's motion is analyzed. The possibility of diagnostics of optical currents in liquids caused by polarization characteristics of an optical field alone is demonstrated using nanoscale metallic particles. We also discuss the prospects of studying temporal coherence using the proposed approach.

There is the motivation of this study. Experimental investigation and computer simulation of the behavior of small spherical particles embedded in optical fields provide a deeper understanding of the role of the Poynting vector for description of optical currents in various media [24, 36]. So, interference between waves polarized in the plane of incidence has been shown to be effective in creation of polarization micro-manipulators and optical tweezers. On the other hand, this is a vital step in optimal metrological investigation of optical currents in vector fields [39 - 42]. Besides, the study of spatial and temporal peculiarities of the motion of particles embedded in optical fields with various spatial configurations and with various scale distributions of the Poynting vector leads to new techniques for estimating the temporal coherence of optical fields [43].

Computation of the spatial distribution of the time-averaged Poynting vector determining the forces affecting microparticles and their movement is performed following the algorithm proposed by M. Berry [24] who has shown that the vector force affecting a small particle in an optical field is proportional to the time-averaged Poynting vector. We will show that the study of the motion of microparticles in inhomogeneously polarized fields provides reconstruction of the spatial distribution of the time-averaged Poynting vectors, *viz.* the optical currents.

### 3.1 Two-wave superposition for changeable degree of mutual coherence of the components

Superposition of two plane waves of equal amplitudes polarized in the plane of incidence (Fig. 2a) results in the distribution of the Poynting vector shown in Fig. 2b. Such distribution arises when the interference angle is equal to  $90^\circ$ , and the only periodical polarization modulation of the field (in the absence of intensity modulation) takes place in the plane of observation [44].

The coherency matrix  $W(\mathbf{r}_1, \mathbf{r}_2, t)$  describes the coherence properties of vector optical fields, being characterising the correlation of two fields at two different spatial points  $\mathbf{r}_1$  and  $\mathbf{r}_2$  [45, 46], and is determined as

$$W(\mathbf{r}_1, \mathbf{r}_2, t) = \langle E_i^{(1)}(\mathbf{r}_1, t) E_j^{(2)*}(\mathbf{r}_2, t) \rangle$$

where  $i, j = x, z$ . Within the framework of such approach the degree of mutual coherence of the field is defined as [15]

$$\eta_{ij}(\mathbf{r}_1, \mathbf{r}_2, t) = \frac{W_{ij}(\mathbf{r}_1, \mathbf{r}_2, t)}{\sqrt{\text{tr}[W(\mathbf{r}_1, \mathbf{r}_1, 0)]} \cdot \sqrt{\text{tr}[W(\mathbf{r}_2, \mathbf{r}_2, 0)]}} = \frac{W_{ij}(\mathbf{r}_1, \mathbf{r}_2, t)}{\sqrt{\sum_{ij} W_{ii}(\mathbf{r}_1, \mathbf{r}_1, 0) W_{jj}(\mathbf{r}_2, \mathbf{r}_2, 0)}} \quad (3)$$



The distribution of the time-averaged density of the energy current in space determines the current magnitude at different points of the plane of observation, being unambiguously determined by the degree of coherence of the superposing waves. The direction of the resulting current is set by the directions of the Poynting vectors of these waves.

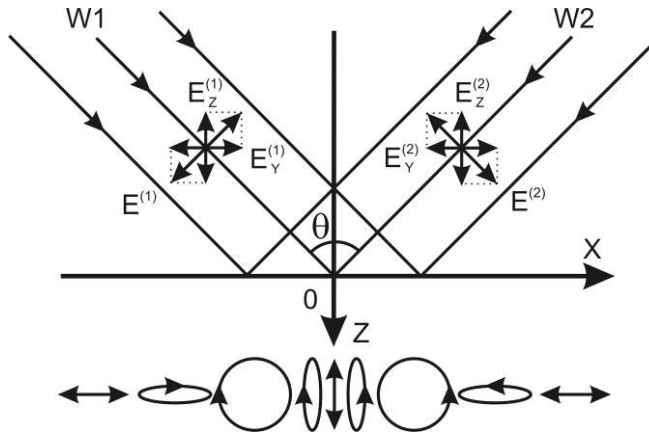


Fig. 2a. Superposition of plane waves of equal amplitudes linearly polarized in the plane of incidence having an interference angle of  $90^\circ$ . Periodical spatial polarization modulation takes place in the plane of incidence.

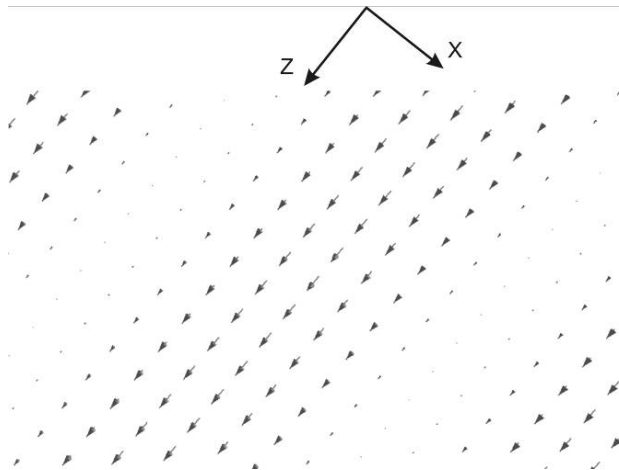


Fig. 2b. Spatial distribution of the time-averaged Poynting vectors resulting from superposition of two orthogonally linearly polarized waves with an interference angle of  $90^\circ$ .

An analysis of the spatial distribution of the time-averaged Poynting vectors shown in Fig. 2b reveals the periodicity of this distribution, where the lengths of lines shown in the figure are proportional to the absolute magnitudes of the vectors. The lines corresponding to the singularities of the Poynting vector are shown by the indicated set of points [47 - 49].

Spatial distribution of the time-averaged Poynting vectors, cf. Fig. 2b, shows the trajectories of energy transfer. The points at the map of the time-averaged Poynting vectors (Fig. 2b) correspond to the areas through which energy transfer is absent, showing: (i) the loci of singularities of the Poynting vector; (ii) the points (with constant intensity) forming the directions along which light energy is non-vanishing, but is conserved; (iii) the points where the vector  $H$  vanishes due to interference, while in this arrangement (90°-superposition of plane waves) superposition of strictly coaxial vectors  $H$  of equal amplitudes associated with two superimposed plane waves takes place.

On the other hand, the instantaneous magnitude of the electric (magnetic) field strength's vector of the resulting distribution, which is formed in the plane of observation, is written as

$$\mathbf{E} = \left| \mathbf{E}^{(1)} + \mathbf{E}^{(2)} \right| \cos(\omega t + \delta_e) \mathbf{a}_e \quad (\text{or } \mathbf{H} = \left| \mathbf{H}^{(1)} + \mathbf{H}^{(2)} \right| \cos(\omega t + \delta_h) \mathbf{a}_h),$$

where,  $\mathbf{a}_e$ ,  $\mathbf{a}_h$  are the unit vectors in the direction of propagation of the electric (magnetic) components for the resulting field in the plane of observation;  $\delta_e$  ( $\delta_h$ ) is the phase difference of the electric (magnetic) field components of superposed waves. In this case, the instantaneous magnitude of the Poynting vector is

$$\mathbf{S}_{inst} = \mathbf{E} \times \mathbf{H} = |\mathbf{E}| \cdot |\mathbf{H}| \cos(\omega t + \delta_e) \cos(\omega t + \delta_h) (\mathbf{a}_e \times \mathbf{a}_h),$$

and the time-averaged magnitude of the Poynting vector is

$$\mathbf{S}_{ave} = \frac{|\mathbf{E}| \cdot |\mathbf{H}|}{2} (\mathbf{a}_e \times \mathbf{a}_h) \cos(\delta_e - \delta_h) = \frac{1}{2} (\mathbf{E} \times \mathbf{H}) \cdot \cos(\delta_e - \delta_h). \quad (4)$$

Because the phase difference of the electric field changes from point to point (polarization modulation), the time-averaged magnitude of the Poynting vector is modulated in space taking the maximum (minimum) at different points of the plane of observation, as it is seen from Eq. (4).

Homogeneous intensity distribution and periodical spatial modulation of the Poynting vector simultaneously realized in the observation region have previously been discussed within the framework of Refs [50, 51]. Spatial polarization modulation at the plane of observation is caused by superposition of the  $E_x$  and  $E_z$  field components with changing the phase difference from point to point, cf. Fig. 2a. A photodetector registers only intensity,  $I = E_x^2 + E_z^2$ . The sum of the squared amplitudes of the electrical field components is constant at the plane of observation, though the state of polarization changes. The Poynting vector is defined, as mentioned above, by the vector product,  $S = E \times H$ . One observes the dependence of the result on the phase relation between vectors  $\mathbf{E}$  and  $\mathbf{H}$  through the vector magnitude and its direction. This relation changes from point to point in the plane of observation and manifests itself in polarization modulation. An obvious explanation for this follows from consideration of the vector product of the components of vector  $\mathbf{E}$  ( $E_x$  and  $E_z$  components) with vector  $\mathbf{H}$ . Both the magnitudes of projections  $E_x$  and  $E_z$  and their phases change from point to point in the observation plane. As a consequence, the vector product changes as well as the Poynting vector. The result of modulation is shown in Fig. 3.

The results of simulating the motion of particles embedded in the field of the considered distribution of the Poynting vector are shown in Fig. 4. We have here tacitly assumed the particles to be absorbing and of size  $0.1 \mu\text{m}$ . One observes that in the case of the distribution resulting from superposition of completely mutually coherent waves, the velocities of particle motion along the lines of maxima and zeroes of the Poynting vector are considerably different from one another.

The particle size is here comparable with a half-period of the corresponding distribution; however, the resultant force giving rise to the particle motion along the lines close to the Poynting vector maxima exceeds the resultant force for lines close to the zeroes of the Poynting vector. The results of modulation of particle movement velocity along the peaks and zeroes of the averaged field of the Poynting vector are shown in Fig. 4a and Fig. 4b, respectively.

If the degree of mutual coherence of the superposed waves equals 0.2, the spatial distribution of the averaged Poynting vectors becomes more homogeneous, the modulation depth decreases considerably, and the velocities of microparticles become almost identical.

When the degree of mutual coherence equals 0.5, the relative velocities of the microparticles along the same trajectories are lower in comparison with velocities in case of complete mutual coherence of the superposed waves and lie in the vicinity of the average magnitudes for coherent and incoherent cases [52]. One observes the dependence on the coherent properties of the superimposed waves for the motion velocities of microparticles with constant size and form in media with constant viscosity [52]. When analyzing the motion of test particles in the region of distributed magnitude of the Poynting vector, the influence of the parameters of superposing fields on the character of particle motion can be determined, cf. Fig. 4a, 4b.

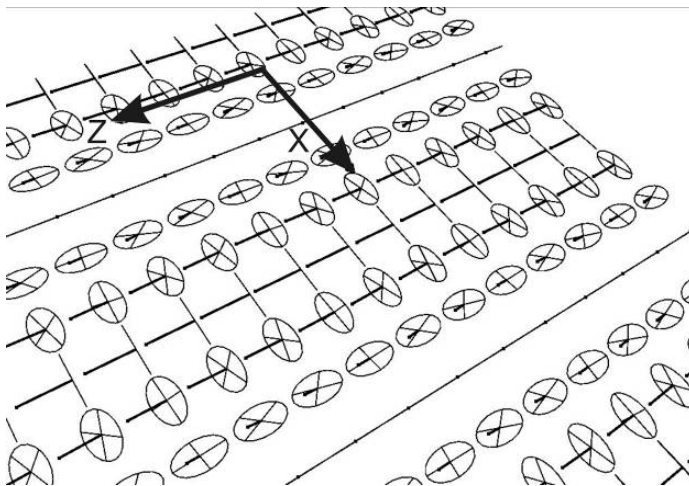


Fig. 3. The polarization distribution in the registration plane is marked by thin lines. The direction and magnitude of the Poynting vector are marked by bold lines. The point at the end of the vector determines the energy transfer direction. The modulation of the Poynting vector takes place according to the polarization modulation at the plane of observation.

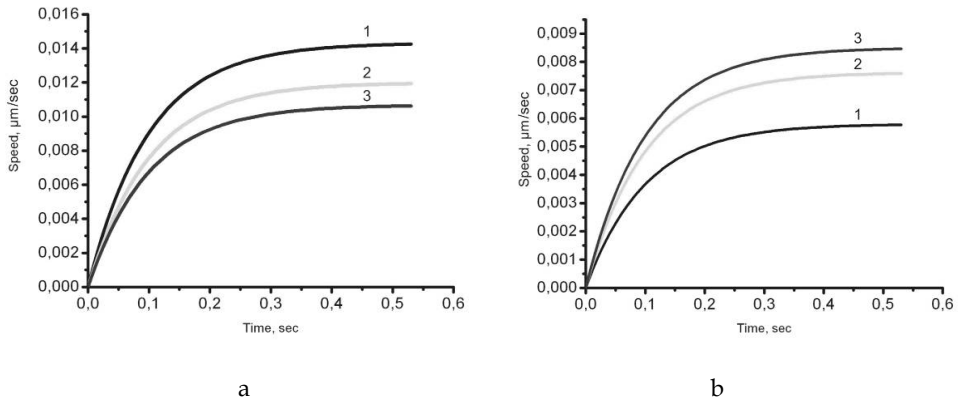


Fig. 4. The change of the particle motion velocity with time obtained for different magnitudes of the degree of coherence of superposing waves in the case of particles moving along the peak (a) and the minimum (b) of the field of time-averaged Poynting vector magnitude: curves 1, 2, and 3 correspond to the degree of coherence, which equals 1, 0.5, and 0.25, respectively.

When the degree of mutual coherence equals 0.5, the relative velocities of the microparticles along the same trajectories are lower in comparison with velocities in case of complete mutual coherence of the superposed waves and lie in the vicinity of the average magnitudes for coherent and incoherent cases [52]. One observes the dependence on the coherent properties of the superimposed waves for the motion velocities of microparticles with constant size and form in media with constant viscosity [52]. When analyzing the motion of test particles in the region of distributed magnitude of the Poynting vector, the influence of the parameters of superposing fields on the character of particle motion can be determined, cf. Fig. 4a, 4b.

As it was shown in papers [44, 51, 52], the degree of coherence of superposing waves determines not only the visibility of an interference distribution, but also the structure of the polarization field, *viz.* it determines the distribution of the Poynting vector. Under the same other conditions of the wave superposition, *changing the degree of coherence results in changing motion velocity of the test particles, what can serve as an estimating parameter for determining the coherence properties of superposing waves.* These differences in velocities of motion of microparticles are explained physically in the following manner: Increasing the share of incoherent radiation in the resulting field distribution causes a decrease of the modulation depth of the Poynting vector's spatial distribution, as well as a decrease of the resultant force magnitude along the lines of energy transfer, which induces the motion of the microparticles. The increase of the degree of coherence brings about an accelerated particle motion in the field of averaged energy magnitudes.

The following diagram (Fig. 5) shows the particle velocity distribution (in this case, 39 particles), embedded into the field formed by the averaged magnitudes of the Poynting vector in the case of superposition of completely mutually coherent waves. With time ( $\sim 1.2$  sec) practically all particles gain equal velocity magnitudes (see column 7); redistribution of particles in the direction of the resultant force and uniform motion along the zero value of the Poynting vector take place.

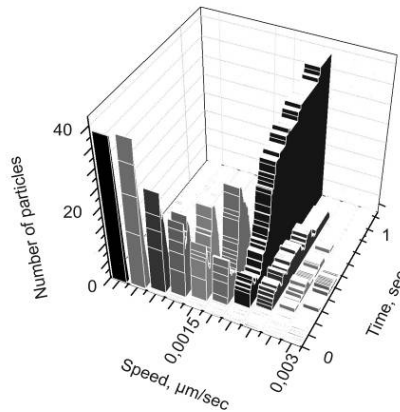


Fig. 5. The diagram of particles velocity distribution with time.

### 3.2 Superposition of four waves for changeable degree of mutual coherence of the components

In the case of superposition of four waves, see Fig. 6a, involving two sets of counter-propagating plane waves of equal intensities, linearly polarized in the plane of incidence and oriented at an angle of  $90^\circ$  with respect to each other, the spatial distribution of the time-averaged Poynting vectors is formed as shown in Fig. 6b.

The 2D periodicity of the Poynting vector's distribution is evident. As in the previous case, the lengths of the time-averaged Poynting vectors are proportional to their magnitudes. The nodal points in this distribution correspond to zero magnitudes of the Poynting vector, i.e. singularities of the Poynting vector. In the following simulation, the diameters of the particles are changed to be comparable with a half-period of the corresponding spatial distribution of the Poynting vector.

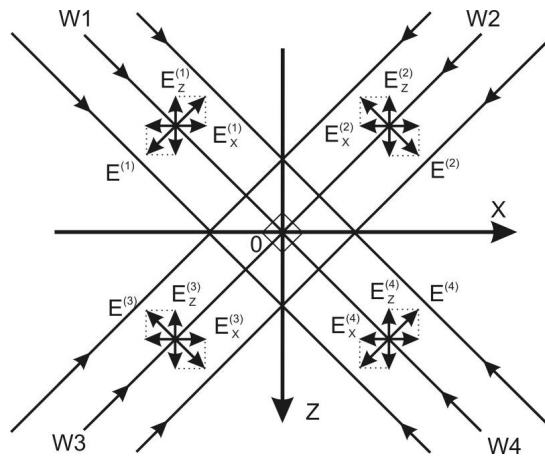


Fig. 6a. Arrangement of superposition of four plane waves.

This follows from the presence of the minimum of the modulation depth at the spatial distribution of the Poynting vector. If the phase relations between four superposed beams are such that the modulation depth of the spatial distribution of the Poynting vector is maximal, the particle velocities will depend on the degree of mutual coherence between the interfering beams, see Fig. 7.

In order to compare the influence of the temporal and spatial parameters of coherence on the motion of the microparticles, we have analyzed the maps of the time-averaged Poynting vector with a superposition of four plane waves over a large area. During this, we have tracked the microparticles' motion. The dependence of microparticles' velocities on the phase difference of the superposing beams has thus been revealed. So, in the case of pair-by-pair four opposite-in-phase superposed beams, particles become motionless. For that, the "opposite-in-phase" configuration covers the situation where two sets of mutually orthogonal standing waves are characterized by the fact that their nodes strictly coincide.

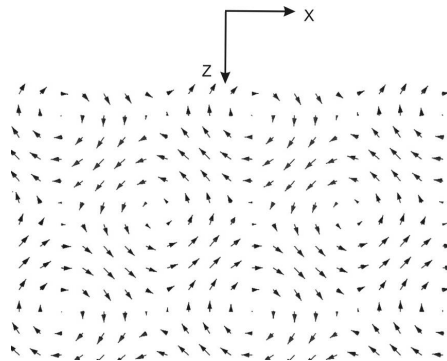


Fig. 6b. 2D distribution of the time-averaged Poynting vectors resulting from the superposition of four waves shown in Figure 6a.

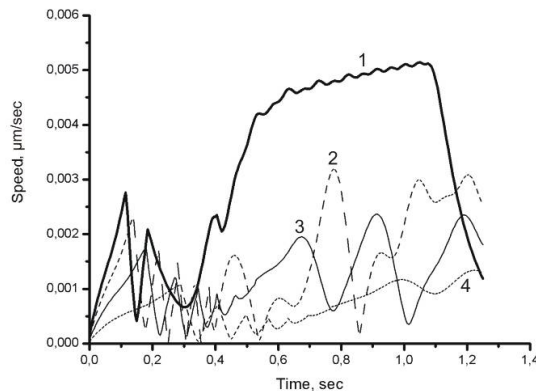


Fig. 7. The variation of motion velocity of a test particles in an averaged field of distributed Poynting vectors with the change of the degree of mutual coherence of the waves (four superposing waves are in phase): curve 1 – one of the waves is incoherent; curves 2, 3, 4 correspond to the degree of coherence 0.25, 0.5, and 0.75, respectively.

Increasing the degree of mutual coherence of the waves sets a more uniform velocity magnitude of moving particles. The magnitude of the resultant force causing this motion under increasing the degree of coherence, practically, does not change with time, see Fig. 8. The maximum depth of modulation for coherent equiphase waves determines the stable position of particles. *The chaotic state and the average particle velocity value can be taken as a possible guideline in estimating the degree of coherence of superposing waves.*

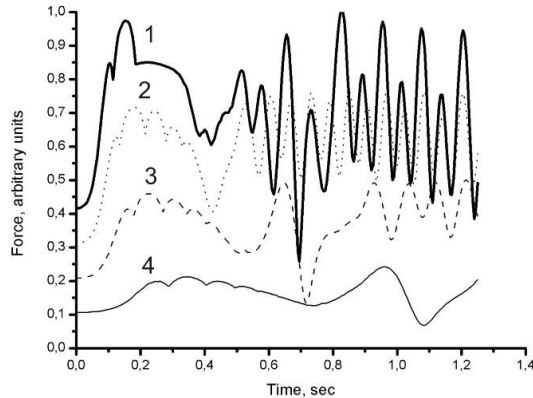


Fig. 8. The change of the resultant force of the test particle motion in the time-averaged field of distributed Poynting vectors with the change of the degree of mutual coherence of the waves (four superposing waves are in phase): curve 1 – one of the waves is incoherent with all other waves; curves 2, 3 and 4 correspond to the degree of coherence of the waves 0.25, 0.5, and 0.75, respectively.

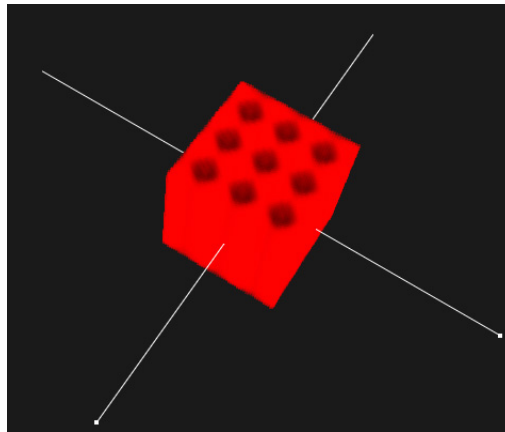


Fig. 9. "Cellular" distribution of the potential traps for microparticles in the case of superposition of four waves.

It is worth emphasizing two issues for the case of superposition of four plane waves. First, the dependence of the depth of modulation for the distribution of the time-averaged

Poynting vectors on the phase relation of superposing waves. It is here assumed that changing the phase relation between the superposed waves causes a transition (in the observed pattern) from the situation when the maxima of two systems of mutually orthogonal standing waves coincide to the case when the nodes of two such systems coincide. Thus, the velocities of particles in such fields depend on the depth of modulation of the distribution of the time-averaged Poynting vector, as it is seen in Fig. 8. Second, the superposition of four waves linearly polarized in the plane of incidence results in forming so-called "cellular" structure in the resulting field distribution, see Fig. 9, which can be used for transfer (transporting) of the set of periodically positioned microparticles *as an entity* to a desired zone.

One considers a future deeper investigation of the peculiarities of motion of microparticles to reveal the coherent characteristics of the waves constituting certain spatial polarization distributions.

The use of strongly reflected test spherical particles provides obtaining more realistic notion on movement of particles in the field modulated in polarization in the incidence plane. So, the test particles are concentrated in zones (planes) of minima of the time-averaged Poynting vector and move along these planes. This situation reflects in the most adequate manner the processes of particle moving in the fields spatially modulated in polarization.

### 3.3 Experimental technique and results

Direct experimental verification of the results of computer simulation is rather difficult. Spatial period of the polarization distribution resulting from superposition of plane waves meeting at right angle is less than a wavelength of the laser radiation of the visible range. In this case, diagnostics of optical currents presumes the using test particles (preferably spherical) of size much less than the period of polarization distribution. That is why, direct visualization and diagnostics of such particle currents is hampered.

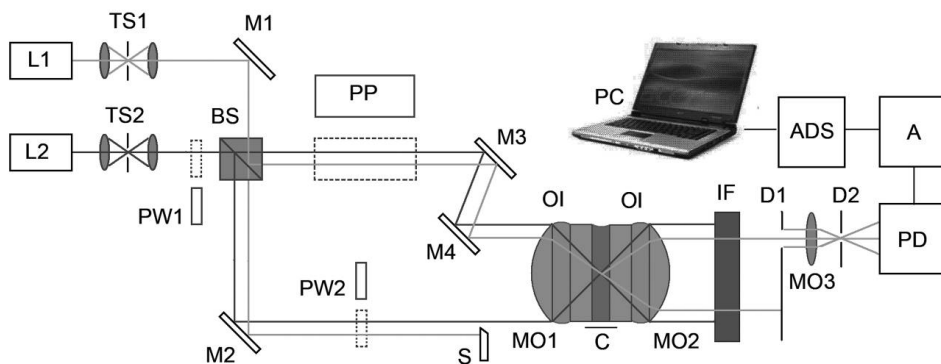


Fig. 10. Experimental setup: L1, L2 - lasers; TS1, TS2 - telescopic systems; M1, M2, M3, M4 - mirrors; PW1, PW2 - half-wave plates for  $\lambda=635$  nm; PP - plane-parallel plate; BS - beam-splitter; MO1, MO2, MO3 - microobjectives; C - cuvette with gold hydrosol; IF - interference filter at  $\lambda=532$  nm; D1(0.7-diam), D2 - diaphragms; S - opaque screen; PD - photodetector; A - amplifier; ADC - analog-to-digit converter; PC - computer.



For verifying the results of above consideration and computer simulation, we have studied experimentally the influence of the field resulting from superposition of two plane waves meeting at right angle with various combinations of their states of polarization on the test particles. Experimental arrangement is shown in Fig. 10. To provide right angle between the beam axes, we use immerse-oil-microobjective  $90\times$  with NA 1.25. Two parallel linearly polarized beams converge at the focus of microobjective 1. If the electrical vectors of two beams are parallel, the intensity distribution as a set of interference maxima and minima is formed at the area of superposition of such beams. We used radiation from a semiconductor laser RLTMRL-III-635 ( $\lambda=635$  nm). For that, the period of an interference pattern is 449 nm. For investigation of the influence of the field distribution with such period on the particles, particle's size must be much less than the mentioned period. We have used spherical particles of hydrosol of gold with diameter 40 nm, approximately. Hydrosol have been obtained following the standard technique [53], by mixing of chloroauric acid ( $\text{H}[\text{AuCl}_4]$ ) and sodium citrate ( $\text{Na}_3\text{Cyt}$ ). Let us describe the experimental conditions in more details. Weight concentration of gold particles in water was  $5 \cdot 10^{-6} \text{kg/m}^3$ . Experiment was carried out at room temperature, humidity 65% and air pressure 741 mm Hg. Note, our observation show that modest changing of the experimental conditions against ones mentioned here do not influence appreciably the results of experiment. As a matter of fact, we perform experiment with low-power of laser radiation (not exceeding 5 mW). In this case, one can neglect acoustic waves arising due to thermal action of laser radiation, which would become very important when one operates with high-power impulse laser beams. On these reasons, special precautions, such as use of anechoic chamber, are not undertaken in our study.

Periodical intensity distribution causes movement of the particles and formation of the periodical distribution of concentration of particles as the planes coinciding with interference minima of the intensity distribution at the area of superposition of two beams. These planes can be regarded as the analog of crystallographic planes in crystals. Direct visualization of particles and their currents is hampered due to small particle size. However, at planes of dense packing of particles self-diffraction takes place. We have observed this phenomenon for angles of meeting of two beams less than  $40^\circ$ . For right angle of meeting of the beams, the each self-diffracted beam propagates *along of* and *contrary to* the propagation direction of other of two superposing beams. Thus, it is impossible to discriminate the initial and self-diffracted beams. That is why, taking into account the Bragg law, we use, for diagnostics of periodical distributions of particles, the test laser beam with another wavelength,  $\lambda=532$  nm. To form the same interference distribution (with the period 449 nm) with such wavelength the angle of meeting of two beams could be  $72.6^\circ$ . So, the angle of incidence of the probing beam must be  $36.3^\circ$  in respect to bisector of the writing beams. In this case, the Bragg law is fulfilled strictly for the probing beam. The mentioned angles are the angles of propagation in light-scattering media, in our case in water.

Two some shifted beams from a green laser (marked by thin line in Fig. 10) propagate in parallel to the beams of red laser (marked by thick line). The external green beam is stopped by the screen S, while the inner probing beam passes the microobjective MO1 and falls at the angle  $36.3^\circ$  into the area of interference extrema. Diameters of the focused beams of red and green lasers are approximately  $12 \mu\text{m}$  and  $10 \mu\text{m}$ , respectively. Glass cuvette C with gold hydrosol is placed at the area of interference pattern. Thickness of the cuvette walls  $0.15$  mm, and thickness of the swept volume is  $18 \mu\text{m}$ . Oil immersion with refraction index 1.515 is placed between microobjectives MO1 and MO2. Microobjective MO2 is used for adjusting the optical arrangement and output the radiation diffracted on periodical distribution of gold particles.

Optical lengths of two legs of an interferometer BS-M2-M3-M4-MO2 are strictly identical. So, two beams from red laser are mutually coherent and interfere at the focus of the microobjective MO2. Placing a perfect plane-parallel plate PP of thickness 19 mm into one interferometer leg leads to disappearance of interference, while the corresponding optical path difference exceeds the coherence length of red laser. What is important, introducing the plate PP must not be accompanied with shifting beams into interferometer. In such a manner, one can control appearance and disappearance of interference extrema at the focus of MO1.

The change of the position of the plate leads to the change of the photodetector signal. Thus, in the case of superposition of radiations from lasers L1 and L2 which are linearly polarized at the plane perpendicular to the figure plane removal of the plate PP results in increasing signal from a photodetector. It shows forming periodical spatial distribution of gold particles and appearance of the diffracted probing beam. The diffracted signal appears for radiation power of red laser more than 2 mW. However, it has been observed that for radiation power exceeding 50 mW, non-linear effects occur in light-scattering medium. So, gold particles absorb radiation and heat environment, acting as a thermal lens. That is why we have carried out our experiment for radiation power of red laser 5 mW. Radiation power of the probing beam was 0.5 mW, so that it can not affect gold particles.

If two beams of red laser are polarized in the figure plane (half-wavelength plate PW1 for  $\lambda=635$  nm is inserted) and their convergence angle is equal to  $90^\circ$ , only polarization modulation takes place in the incidence plane. In this case the diffracted probing beam is present as well. The signal at the photodetector output with and without plane-parallel plate is shown in Fig. 11 b. The diffracted probing beam is present, but is approximately of half the intensity in comparison with the case illustrated in Fig. 11 a. This experimental result is also in accordance with the result of computer simulation. The spatially modulated in polarization field (in the plane of incidence) is correlated with concentration of the test particles at the planes of minima of the time-averaged magnitude of the Poynting vector, and particles move along these planes.

If two beams from red laser are linearly polarized, but one of them in the figure plane while another one perpendicularly to this plane (a half-wave plate PW2 for  $\lambda=635$  nm is inserted), the diffracted probing beam is absent, cf. Fig. 11 c. This shows that at the focal plane where the beams from red laser superpose, the periodical distributions of gold particles are absent. This experimental result is also in agreement with earlier computer simulation [42]. In other words, there are no any ordered optical currents being liable to optical diagnostics, as it has been made in previous case.

Thus, temporal and space peculiarities of particle's motion in optical fields without intensity modulation, but only due to polarization modulation causing the spatial modulation of the time-averaged Poynting vector (depending on the degree of mutual coherence of superpose waves) opens up new feasibilities for the use of such field characteristics and the parameters of microparticles motion for estimating the temporal coherence of the tested field. Here we have demonstrated a possibility of influence of only the polarization factor on formation of optical currents in liquids by the use of the principles of spatial polarization modulation in the observation plane. Besides, we have shown the possibility of diagnostics of optical currents using test particles of nanoscale.

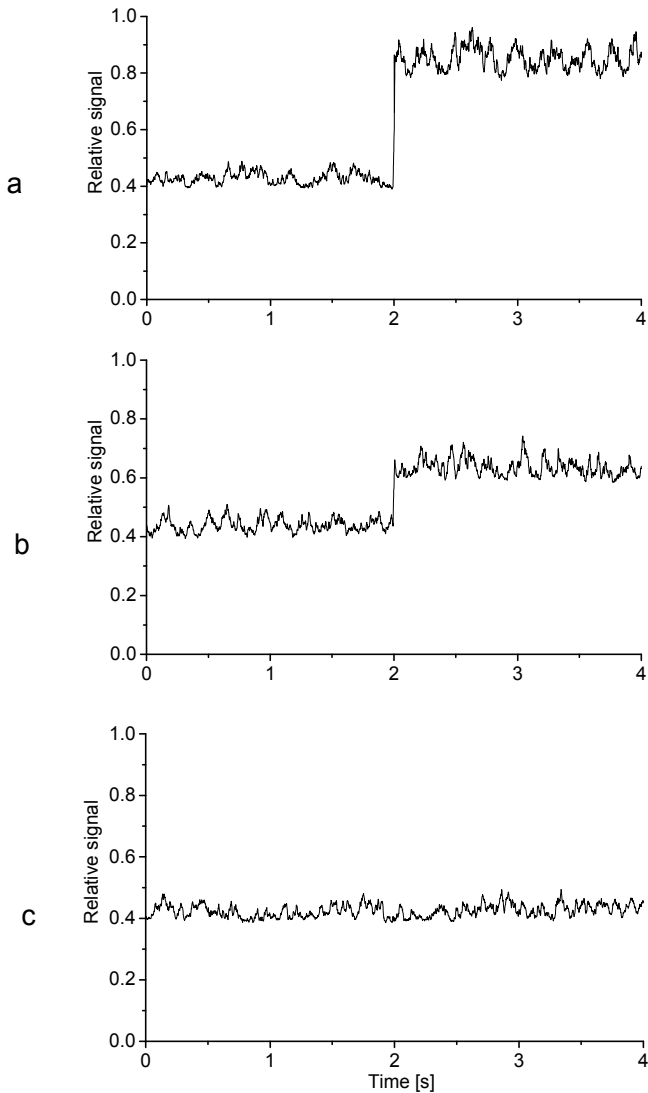


Fig. 11. Relative signal of a photodetector (a plate PP is inserted on 2 sec and then removed) in the case when radiation of red laser is linearly polarized: (a) both beams are polarized in the plane perpendicular to the figure plane; (b) both beams are polarized in the figure plane; (c) one beam is polarized at the figure plane, while another one is polarized perpendicularly to this plane.

The explained metrology of microstructure of optical fields may be extended, to all appearance, on polychromatic waves. The initial steps in this direction have been recently made in studies [52, 54 – 60].

#### 4. Polarization singularities in partially coherent light beams

In this section we describe specific polarization singularities arising in incoherent superposition of coaxial orthogonally polarized laser beams. It is shown that in transversal cross-section of paraxial combined optical beams of this class, instead of common singularities, such as amplitude zeroes (optical vortices) inherent in scalar fields [2], and polarization singularities such as C points and L lines inherent in completely coherent vector fields [18], *phase singularities of the complex degree of polarization* (CDP) arise, whose description and investigation have been initiated by papers [61 - 65] basing on earlier studies [66 - 68] concerned to the Young's concept of the edge diffraction wave in connection with diagnostics of phase singularities of spatial correlation functions of optical fields. There are U contours along which the degree of polarization equals zero and the state of polarization is undetermined (singular), and isolated P points where the degree of polarization equals unity and the state of polarization is determined by the non-vanishing component of the combined beam. (Note, discussed here notion of CDP differs from the definition of the complex degree of mutual polarization, CDMP [6] that is two-point function of an optical beam.)

Let us briefly argue the relevance of the introduced approach.

It is known [69] that each level of description of optical phenomena possesses its own set of singularities, *i.e.* the set of elements of a field (points, lines, surfaces, depending on considered dimension) where some parameter of a field is undetermined. Importance of detecting such elements of a field is caused by the fact that such elements form peculiar skeleton of a field, so that if one knows behavior of a field at such singular elements (and at nearest vicinities of them), one just can predict, at least in qualitative manner, but with high level of validity, behavior of a field at all other areas.

Conditionally, one can classify singularities of optical fields in the following manner [69]:

- singularities of geometrical optics,
- singularities of completely coherent scalar (homogeneously polarized) wave fields – optical vortices,
- singularities of completely coherent vector (inhomogeneously polarized) wave fields – optical vortices,
- singularities of partially coherent wave fields,
- singularities of quantum optics – “quantum cores”.

Singularities of geometrical optics are caustics where the field amplitude reaches infinity. Singularities of completely coherent wave fields are divided into two sub-classes: (i) for scalar (homogeneously polarized) fields and (ii) for vector (inhomogeneously polarized) fields. In scalar fields, when polarization can be neglected, so-called wave front dislocations take place (which are also referred to as amplitude zeroes or optical vortices). Phase of the complex amplitude is undetermined at such elements and is step-like changed at crossing of them. In vector fields optical vortices are absent, though they remain in any polarization (“scalar”) component. Instead of vortices, polarization singularities arise at cross-section of a field, *viz.* field elements where azimuth of polarization (C points) or handedness (L lines) is

undetermined [18]. Vector skeletons of coherent inhomogeneously polarized fields were elaborated in details in papers [70 - 72]. By crossing L lines, handedness is step-like changed into opposite one; by crossing C point, azimuth of polarization is changed into orthogonal one. These types of singularities are blurred in quantum-mechanical description being "camouflaged" by so-called quantum vacuum [69], though the distance of influence of such quantum core is rather small. Its linear size is of order of magnitude  $\lambda^{-3}$  (about 6 Å for He-Ne laser).

All mentioned singularities disappear in the case of partially coherent wave fields (though they remain in each completely coherent component, mode in a set of which partially coherent radiation is decomposed. Instead (beside) of them, new singularities appear inherent just in partially coherent fields. Let us emphasize, that singularities of partially coherent fields have formed the novel topic in the field of singular optics just at the beginning of the Third Millennium [3].

For that, two situations arise again: (i) scalar case when polarization can be ignored while the state of polarization is the same at all point of a field, and (ii) vector case when the state of polarization of *partially coherent* field changes from point to point that requires explicit taking into account of vector nature of light. The first (scalar) case became the subject of intense investigations in last years [67, 73 - 76]. As a result of these investigations, new phase singularities of spatial and temporal correlation functions of quasi-monochromatic light fields have been revealed, as well as singularities of spectral components of polychromatic ("white-light") radiation [68, 77 - 82]. For that, vector singularities of partially coherent light fields have been revealed just recently [61 - 65]. Such singularities are elaborated in this section.

#### 4.1 U and P singularities in partially spatially coherent combined beams

Let us consider vector singularities in partially coherent optical beams by giving the following simple instructive example. Mutually incoherent and orthogonally polarized Laguerre-Gaussian mode LG01 and a plane wave are coaxially mixed. Such components can be obtained from one laser (using a computer-generated hologram for forming LG01 mode) in interferometric arrangement with optical delay,  $\Delta l$ , considerably exceeding a coherence length of the used laser,  $l$ , or using two different lasers. Intensity of a plane wave is set deliberately to be less than the peak intensity of the mode, see Fig. 12.

Thus, we consider two-component mixture co-directional orthogonally polarized beams, one of which contains a common phase singularity, *viz.* optical vortex. Interference between such beams with forming common interference fringes is excluded by two reasons: (i) specified mutual incoherence of the components: (ii) polarization orthogonality of them. Note, even only the second condition *per se* determines that, independently on the degree of mutual coherence of two beams over whole interval from zero (for optical path difference exceeding the coherence length) to unity (for zero optical path difference) *visually* observed and *photometrically* measured pattern remains unchangeable. However, more delicate polarization analysis of the combined beam enables to differentiate two limiting cases, *viz.* completely coherent and completely incoherent mixing of orthogonally polarized components.

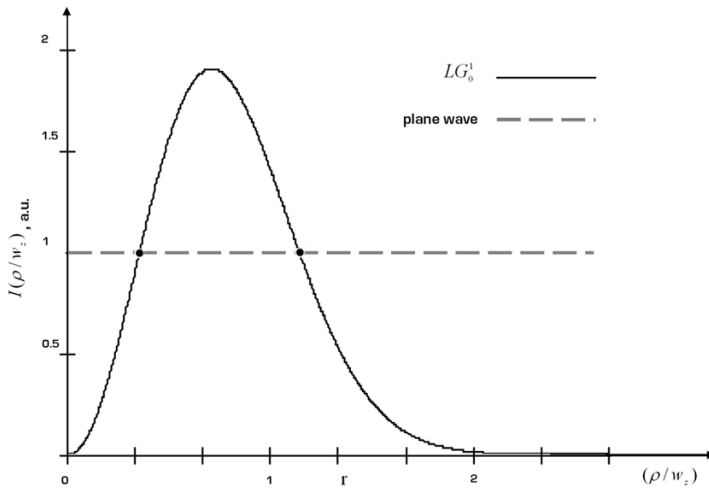


Fig. 12. Mixing of vortex-supporting LG01 mode and plane wave with intensity less than peak intensity of a mode.

Let us firstly consider the limiting case when two components are completely mutually coherent. For the sake of distinctness (and for substantiveness of further consideration), we consider coherent mixing of orthogonally *circularly* polarized LG01 mode and a plane wave. Beside of all, choice of circular polarization basis possesses the advantage that it is invariant in respect to rotation of the coordinates, in contrast to linear or elliptical basis, which are relative ones [83].

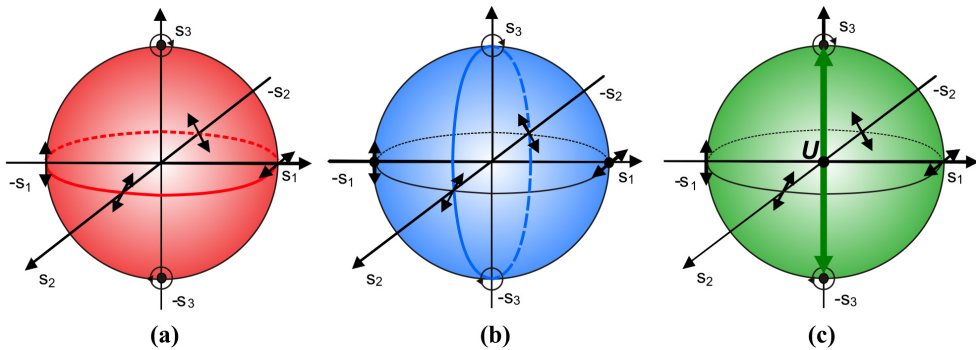


Fig. 13. The lines of equal intensities of orthogonally polarized beams at Stokes space: equator of the Poincare sphere for circular polarization basis, coherent mixing (a); 45°-meridian including the poles for linear polarization basis, coherent mixing (b); diameter of the Poincare sphere connecting the poles for circular polarization basis, incoherent mixing (c).



Let us support this intuitive consideration by formal description. Let us proceed from Jones vectors of two components, right-circularly polarized LG01 mode and left-circularly polarized plane wave,

$$\mathbf{E}_{LG} = c(w/\rho)\exp(i\Delta)\begin{bmatrix} \exp(i\varphi) \\ \exp[i(\varphi + \pi/2)] \end{bmatrix}, \quad \mathbf{E}_P = \begin{bmatrix} \exp(i\varphi) \\ \exp[i(\varphi - \pi/2)] \end{bmatrix}, \quad (5)$$

where  $c$  is the amplitude factor corresponding to inhomogeneous amplitude distribution of a mode as a function of dimensionless radial coordinate, and  $e^{i\Delta}$  is associated with helicoidal change of a phase of a mode under circumference of the central vortex (its explicit form for Laguerre-Gaussian mode is well known but is not relevant here). There is Jones vector of the combined beam:

$$\mathbf{E}_{Total} = \mathbf{E}_{LG} + \mathbf{E}_P = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} c\exp(i\Delta) + 1 \\ c\exp[i(\Delta + \pi/2)] + \exp(-i\pi/2) \end{bmatrix} \exp(i\varphi) \quad (6)$$

General coherency matrix of the beam is found as

$$\{\mathbf{J}\} = \mathbf{E}_{Total} \cdot \mathbf{E}_{Total}^* = \begin{bmatrix} E_x \\ E_y \end{bmatrix} \begin{bmatrix} E_x^* & E_y^* \end{bmatrix} = \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix}, \quad (7)$$

or in explicit form:

$$\{\mathbf{J}\} = \begin{bmatrix} c^2 + 2c\cos\Delta + 1 & c^2\exp(-i\pi/2) + 2c\sin\Delta + \exp(i\pi/2) \\ c^2\exp(i\pi/2) + 2\sin\Delta + \exp(-i\pi/2) & c^2 - 2c\cos\Delta + 1 \end{bmatrix} \quad (8)$$

Combining the elements of coherency matrix, one can find *full* Stokes parameters:

$$S_0 = J_{xx} + J_{yy} = 2(c^2 + 1); \quad S_1 = J_{xx} - J_{yy} = 4c\cos\Delta; \quad (9)$$

$$S_2 = J_{xy} + J_{yx} = 4c\sin\Delta; \quad S_3 = i(J_{xy} - J_{yx}) = 2(c^2 - 1).$$

Here we are especially interested in the case when  $c = 1$ . One just obtains for this case the *normalized* Stokes parameters:

$$s_0 = 1; \quad s_1 = \cos\Delta; \quad s_2 = \sin\Delta; \quad s_3 = 0. \quad (10)$$

Vanishing of the fourth Stokes parameter means that polarization at all points of the contour where intensities of the mixed components are equal to each other are equally distanced from the states of polarization of the components, i.e. neither right-circular nor left-circular predominate in intensity. It is in direct correspondence with Fig. 13 a. At all points of such L contour polarization is linear with the polarization azimuth  $\alpha = 0.5 \tan^{-1}(s_2/s_1) = \Delta/2$ , while the angle of ellipticity  $\beta = 0.5 \arcsin s_3 = 0$ . Besides, the degree of polarization  $P = \sqrt{s_1^2 + s_2^2} = 1$ . In correspondence with helicoidal structure of a wave front of LG01 mode,



a phase difference of the components changes along the contour of equal intensities that results in changing azimuth of polarization. Thus, we obtain direct analog of L contour. Further, at the center of vortex of LG01 mode we have  $c = 0$ . Again, proceeding from Eq. (9) we find the normalized Stokes parameters  $1, 0, 0, -1$ , i.e. left-circular polarization of a plane wave. In the vicinity of such C point polarization is elliptical, with the azimuth of polarization changing with azimuthal coordinate and ellipticity decreasing from the vortex to L contour, Eq. (10), where handedness is undetermined an step-like changing by crossing this contour. It is all in quite correspondence with Fig. 14.

Thus, for circular polarization basis walking along contour of the combined beam "LG01 mode + plane wave" where intensities of the components become equal corresponds to moving along equator of the Poincare sphere that is determined only by the ratio of the second and third Stokes parameters. (For comparison, using linear polarization basis, to say  $0^\circ$  and  $90^\circ$ , one obtains by the same way the normalized Stokes parameters for the combined beam  $1, 0, \cos[\Delta + (\varphi_0 - \varphi_{90})], \sin[\Delta + (\varphi_0 - \varphi_{90})]$  that corresponds to points of  $45^\circ$ -meridian of the Poincare sphere, see Fig. 13 b.)

Before consideration of the most general case of partial mutual coherence of the mixed orthogonally polarized components in the following section, let us consider other limiting case, *viz.* completely incoherent mixing of such components. There is no necessity to proceed now from Jones vectors and to form a coherency matrix of the combined beam. One can at once determine the Stokes parameters of mutually incoherent components and to sum directly them, without accounting phase relations that are irrelevant for incoherent summation. The normalized Stokes parameters of orthogonally polarized beams differ only in sign of the second, third and fourth parameters:  $\{1, s_1, s_2, s_3\}$  and  $\{1, -s_1, -s_2, -s_3\}$ . It is clear that when two components become equal in intensities, the normalized Stokes parameters of the combined beam becomes  $\{1, 0, 0, 0\}$ . The field at such elements of a field is completely unpolarized. There are just U singularities [61 - 63]. This case is shown in Fig. 13 c for the case of incoherent mixing of orthogonally circularly polarizer components. Trajectory of the imaging point for the combined beam in this case is the diameter of the Poincare sphere connecting two poles. U singularity is imaged by the center of this sphere, and all other points (beside the center and poles) image partially circularly polarized fields. For that, the length of a vector drawn from the center of the Poincare sphere to the imaging point inside it equals the degree of polarization. The point where the degree of polarization equals unity is referred to as P (completely polarized) point [61 - 63]. Its location is determined by the vortex of orthogonally polarized (scalar singular) component. The set of P points and U contours corresponding to extrema of the degree of polarization of a field are *the singularities of the degree of polarization* forming the vector skeleton of two-component mixture of orthogonally polarized beams. Note, in papers [61 - 64] consideration is carried out using the notion of the complex degree of polarization – CDP, associated with orientation of the vector of polarization in the Stokes space and undergoing the phase singularity at the center of this space. So, U singularities can be considered just as vector singularities, *viz.* singularities of the vector of polarization, when its magnitude equals zero and a phase (orientation of the vector) is undetermined.

Let us emphasize that the condition of occurring U singularity is equal to the condition of occurring L contour in completely coherent limit. It means that loci of C and L singularities in completely coherent fields and P and U singularities in partially coherent fields arising from completely incoherent orthogonally polarized components, correspondingly, coincide.

Displacement from U singularity results in predomination of one of two orthogonal components in intensity. The state of (partial) polarization is just determined by the predominant component. That is why, the degree of polarization can be determined in similar form as visibility:

$$P = \left| \frac{I_1 - I_2}{I_1 + I_2} \right|. \quad (11)$$

In other words, at each point of the combined beam equal in intensities parts of orthogonal components form unpolarized background, at which manifests itself completely polarized part corresponding to predominant in intensity component. This is in a complete agreement with classical decomposition of partially polarized beam into completely coherent and completely incoherent parts, which are added on intensities, without accounting phase relations [17, 84]. Note, there are no any device providing such decomposition in practice. However, share of completely polarized part can be determined experimentally through the Stokes polarimetric experiment,  $P = \sqrt{s_1^2 + s_2^2 + s_3^2}$  or, equivalently, following Eq. (11). In theory one put in correspondence to such beams the set of two coherency matrices – for completely polarized and completely unpolarized parts of a beam [1, 17].

Thus, only two orthogonal states of polarization take place in combined beams of considered kind, which are separated by U singularities where the state of polarization is undetermined.

So, the considered limiting cases show the same location of C and P singularities and L and U singularities for the same set of components. However, vicinities of such singularities are essentially different. Only two orthogonal states of polarization are present in spatially partially coherent combined beams, and only the degree of polarization changes from point to point within the areas separated by U singularities.

#### 4.2 Vector singularities for partially mutually coherent mixed components

Let us consider now the most general case, when two mixed components shown in Fig. 12 are orthogonally (circularly) polarized and are partially mutually coherent, so that the degree of mutual coherence of the components can be gradually changed from unity to zero. It can be implemented in the arrangement of the Mach-Zehnder interferometer with controllable optical pass difference between the legs of an interferometer, as shown in Fig. 15.

A half-wave plate at the interferometer input serves for fine balancing of intensity ratio between the legs of an interferometer without changing total intensity at its output. Two polarizers inside an interferometer are controllers setting orthogonal linear polarizations. LG01 mode is generated by a computer synthesized hologram. A quarter-wave plate at the output of an interferometer transforms orthogonal linear polarization into orthogonal circular ones.

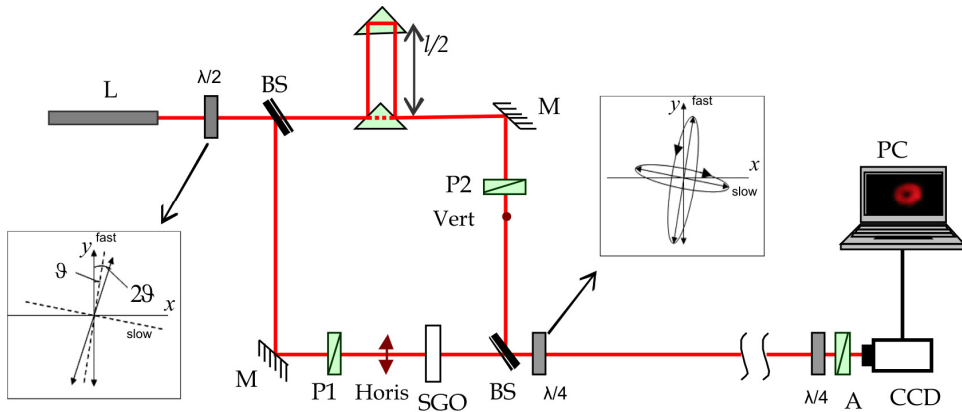


Fig. 15. General arrangement for generation and detection of vector singularities in partially coherent beams: L - He-Ne laser;  $\lambda/2$  and  $\lambda/4$  - half-wave and quarter-wave plates, respectively; P - polarizers; A - linear analyzer; SGO - singularity generating object (computer synthesized hologram); BS - beam splitters; M - mirrors; CCD - CCD-camera; PC - personal computer. Insets show action  $\lambda/2$  and  $\lambda/4$  plates at the input and at the output of an interferometer; two prisms form optical path delay loop.

A quarter-wave plate and linear analyzer at the receiving end, together with CCD-camera matched with personal computer serve for Stokes-polarimetric analysis of combined beams. Two prisms at one of legs of an interferometer enable to control optical path difference and mutual coherence of the mixed components. Namely, one can control path delay  $\Delta l$  from zero to magnitude exceeding a coherence length (length of wave train)  $l$  of the used laser. Change of the ratio  $\Delta l/l$  corresponds to change degree of mutual coherence of orthogonally polarized components. Thus, for  $0 < \Delta l/l < 1$  the combined beam is *simultaneously* partially spatially coherent (due to changing intensity ratio at cross-section of the resulting field) and partially temporally coherent (due to non-zero optical path difference between the components), one expects for increasing optical path difference the following.

As it has been mentioned above, the condition of arising of L contours and U contours in the limiting cases of mixing of orthogonally circularly polarized beams is the same: intensities of the components must be equal to each other. If the optical path difference increases from zero, field at the L contour remains linearly polarized, but the degree of polarization decreases. It follows from that the degree of polarization of a beam is determined by the degree of mutual coherence of its arbitrary orthogonal components [1, 17], here right-hand and left-hand circular components. It means that *U contour nucleates just at the bottom of L contour*.

The degree of polarization can be represented equivalently in terms of *measured* Stokes parameters (that will be used in the next section) or *theoretically*, viz. through the invariants

of the coherency matrix, which at the same time determine coherence properties of a field [17]:

$$P = \sqrt{1 - \frac{4 \det\{\mathbf{J}\}}{\text{Sp}^2\{\mathbf{J}\}}}. \quad (12)$$

For that, in general, the degree of polarization is always non less than modulo of the degree of mutual coherence of the components, for circularly polarized components

$$|\mu_{rl}| = \left| \frac{J_{rl}}{\sqrt{J_{rr}J_{ll}}} \right|. \quad (13)$$

In general case,  $P \neq |\mu_{rl}|$ , as the degree of coherence depends on the decomposition basis while the degree of polarization is invariant [5]. However, it has been shown [1, 17] that the degree of polarization is equal to the *maximal* degree of coherence,  $P \equiv |\mu_{rl}|_{\max}$ , in the case when the components are of equal intensities. This is just the case of L singularities and U singularities. It is of the most importance, that change of the optical path difference changes weights  $|\mu_{rl}|$  of completely coherent (and completely polarized) part of the combined beam and  $1 - |\mu_{rl}|$  of its completely incoherent part. Increasing  $\Delta/l$  difference corresponds to increasing of weight of U singularity against L singularity, so that one can follow gradual transformation of L contour into U contour.

### 4.3 Experimental reconstruction of “pure” and “mixed” polarization singularities

Mixing of LG01 mode and plane wave was performed in the arrangement Fig. 15 [66]. Intensity of a plane wave was considerably (approximately by the order of magnitude) less the peak intensity of a mode, both circularly (orthogonally) polarized. The following results have been obtained under such conditions.

Fig. 16 shows the combined beam whose view, as was mentioned above, within experimental accuracy remains the same at arbitrary optical path delay set in the interferometer. This photo has been obtained for incoherent mixing of two components for  $\Delta/l \approx 3$  (under condition realized in paper [68]. We measured spatial distribution of the Stokes parameters and looked for the elements where  $s_1 = s_2 = s_3 = 0$  ( $P = 0$ , U contours), and  $s_3 = 1$  (P point), see discussion after Eq. (10). In such a manner, we were in a position to reconstruct a vector skeleton of partially spatially coherent combined beam formed by completely mutually incoherent components, as in Refs [61 - 64]. Experimental error in determining the normalized Stokes parameters was at the level 7%; this determines reliability with which we reconstructed P point and U contours. P and U singularities for this case are shown in right fragment of Fig. 16. Two U contours separate the areas with right-circular and left-circular polarization shown by different colors. Within these areas  $\sqrt{s_1^2 + s_2^2} = 0$ , while  $s_3 < 1$ . Let us emphasize that the full Stokes-polarimetric experiment over combined beam cross-section is necessary in this case, as operating only with rotating linear analyzer does not provide differentiation partial circular polarization from complete elliptical polarization.

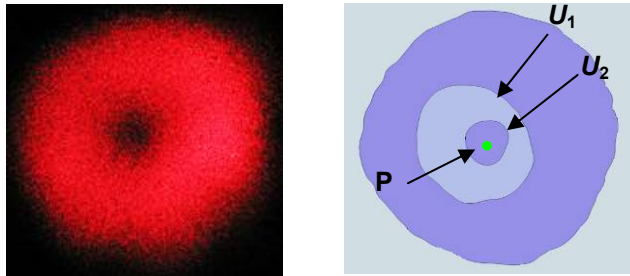


Fig. 16. The partially coherent combined beam (left) and its vector skeleton formed by P and U singularities (right) for completely incoherent mixing of circularly polarized components.

Separate maps of the Stokes parameters are less representative being only raw material for finding out the degree of polarization, ellipsometric parameters of a field, and vector singularities. That is why, we demonstrate separately from 2D pattern shown in right fragment of Fig. 16 1D cross-section of the degree of polarization of this combined beam, see Fig. 17. Red curve shows two-lateral radial dependence of  $P$  computed following Eq. (11). Blue curve shows experimentally obtained distribution found as the combination of measured Stokes parameters, here  $P = |s_3|$ . Quantitative discrepancy of two curves (both in positions of zeroes and in heights of side-lobes) is obvious and is explained by anisotropy of the vortex. Nevertheless, behavior of the experimental dependence is in quite satisfactory qualitative agreement with the simulation results. Namely, one observes two zeroes of the degree of polarization at the each side of the central optical vortex that are the signs of two U contours. Moreover, experiment has proved typical conical vicinity of U contours recently predicted and observed in paper [61], which are reliable sign of true singularity of any kind, in contrast to local minimum.

Another limiting case (completely mutually coherent components) for  $\Delta/l \ll 1$  (approximately 0.05) is illustrated in Fig. 18. Again, spatial maps of the Stokes parameters were obtained and the elements  $s_3 = 0$  and  $\sqrt{s_1^2 + s_2^2} = 1$  were selected. There are the lines of linear polarization. Then, in several selected points of such L lines we determined the azimuth of polarization, again, by two ways: firstly as  $\tan^{-1}(s_2/s_1)$ , and, secondary, as direct measurement of the azimuth of polarization by rotating a linear analyzer up to complete extinction of a field at the specified point that corresponds to crossed azimuth of polarization of the combined beam and the axis of maximal transmittance of analyzer. Description between two results for determining of the azimuth of polarization do not exceeded 0.1 rad.

Perfect extinction of a beam at the specified points just shows that the degree of polarization  $P = 1$  (in contrast to the previous case of completely mutually incoherent components, where intensity at the analyzer output is independent on its orientation). Also, for certain orientations of a quarter-wave plate and analyzer, the field at each other point can be extinguished that shows that everywhere the degree of (elliptical) polarization equals unity. It is worth to compare Fig. 18 with a view of the circular complex polarization plane (Fig. 14) to be convincing of that, really, such polarization distribution over of a combined beam is close experimental analogue of the circular polarization plane.

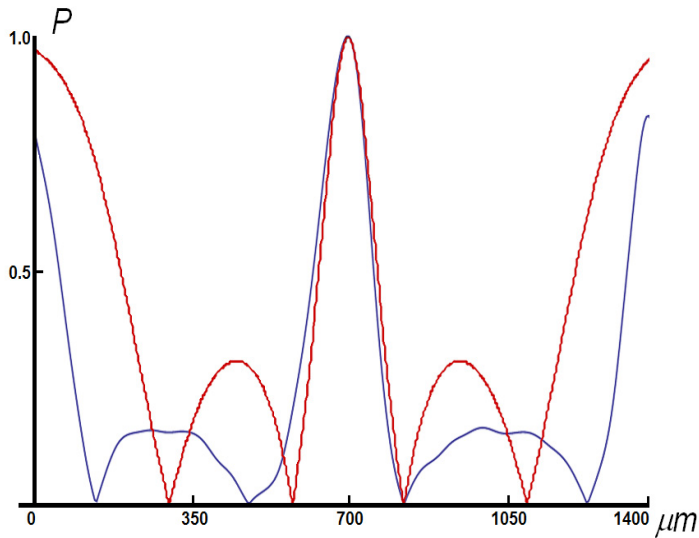


Fig. 17. 1D distribution of the degree of polarization of the combined beam formed by two mutually incoherent orthogonally polarized components defined in Fig.12 and shown in Fig. 16.

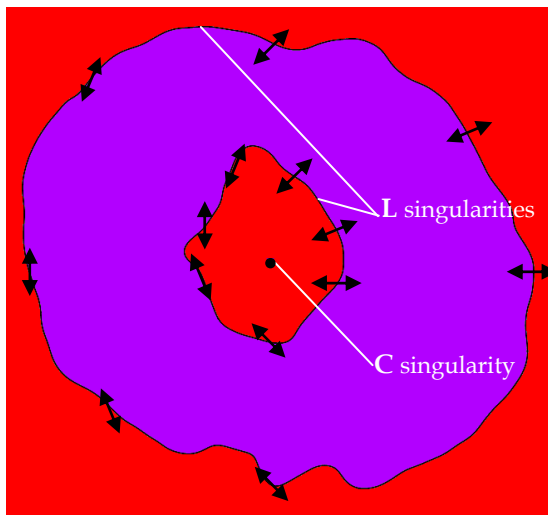


Fig. 18. C and L singularities in combined beam assembled from completely mutual coherent orthogonally (circularly) polarized LG01 mode and plane wave. At L lines, where intensities of two mixed component are equal, the azimuth of polarization changes in agreement of prediction illustrated in Fig. 14 Areas of different colors correspond to opposite handedness.

At last, we have elaborated experimentally intermediate case, when  $0 < \Delta l/l < 1$ , lying between ones considered above. For step-by-step increasing optical path difference between the same orthogonally (circularly) polarized components, we obtained spatial distributions for the Stokes components  $I_0, I_{90}, I_{+45}, I_{-45}, I_r, I_l$  and found from them the Stokes parameters. Further, the degree of polarization and ellipsometric parameters of the combined beam were determined as the combinations of these parameters.

Before formulating the conclusions from our observations, let us represent one of row (intermediate) results undergoing following processing. Fig. 19 illustrates combined beams "LG01 mode + plane wave" (with large intensity ratio, so that one does not visualizes a plane wave) for relative optical path differences close to unity (coherent limit illustrated in Fig. 18) and slightly exceeding a half of the coherence length of used laser, left fragments of Fig. 19. Other fragments of this figure are the intensity distributions  $I_{+45}$  (central column) and  $I_{-45}$  (right column) used for forming the third Stokes parameters. (Other pairs of intensity distributions show the same tendency). Though two orthogonally polarized components do not interfere, their equal polarization projections selected by properly oriented polarizer *can* interfere depending on their mutual coherence. If the degree of mutual coherence of the components is not zero, their equally polarized projections interfere with forming typical patterns indicating phase singularity. To have enough spatial resolution for determining place of vortex, we set non-zero interference angle between the components (which as small enough to be no influencing on accuracy of polarization measurements). For that, instead of snail-like pattern typical for coaxial mixing of LG01 mode and a plane wave, we obtained interference forklets. Comparison of the central and left columns of Fig. 19 shows that spatial intensity distributions for orthogonal polarization projection of the combined beam are complementary in a sense that dark forklet is replaced by bright one.

The main conclusion follows from comparison of fragments b and e (c and f). Decreasing of the mutual coherence of the mixed components and decreasing of the degree of polarization of the combined beam are accompanied by decreasing of ability of equal polarization projections of the mixed components to interfere that manifests itself in decreasing of visibility of interference pattern. So, in fragments b and c of Fig. 19 ( $\Delta l/l \approx 0.05$ ) the measured visibility is 0.97, while in the fragments Fig. 19 e and f ( $\Delta l/l \approx 0.56$ ) visibility is 0.24 (with experimental error non exceeding 5%). It shows the feasibility allows determine *the degree of mutual coherence* of two orthogonally polarized beams by measuring *the degree of polarization* of the combined beam formed by such components found from Stokes parameters. Namely, in our experiment  $|\mu_{ri}|$  for  $\Delta l/l \approx 0.56$  also equals 0.24. For the reasons discussed in Section 3, such measurements are preferably be performed at the elements of the combined beam where intensities of two beams are equal to each other (where L and U singularities co-exist in case of partial mutual coherence of the components), while at such singular elements of the combined beam  $P \equiv |\mu_{ri}|$ .

Thus, vector singularities occurring in light fields, which are simultaneously partially spatially and partially temporally coherent have been considered. It has been shown that in the case of partially coherent mixing of two orthogonally *circularly* polarized components conventional vector singularities, *viz.* C points and L lines submerged in a field of elliptical

polarizations coexist with singularities arising just in partially coherent fields, such as U and P singularities as the extrema of the degree of polarization. Gradual transformation of C and L singularities into P and U singularities, respectively, accompanying decreasing degree of mutual coherence of the components has been experimentally shown. So, conventional polarization singularities of completely coherent fields (C points and L lines) are vanish in incoherent part of the combined beam, so that the only polarization of the component predominant in intensity remains in the vicinities of P points and U lines.

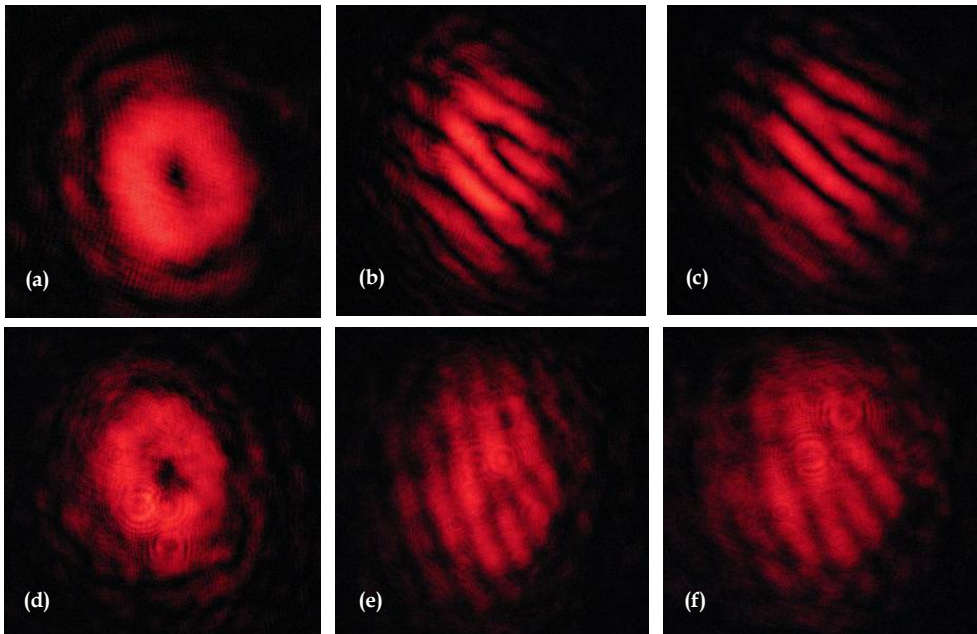


Fig. 19. The combined beams “LG01 mode + plane wave” with relative optical path differences  $\Delta l/l \approx 0.05$  (a) and  $\Delta l/l \approx 0.56$  (d); the corresponding intensity distributions behind a linear analyzer for determining the third Stokes parameters:  $+45^\circ$  (b and e) and  $-45^\circ$  (c and f). Decreasing visibility of interference fringes in fragments e and f corresponds to decreasing in parallel the degree of mutual coherence of the mixed components and the degree of polarization of the combined beam.

## 5. Feasibilities for experimental analysis of characteristics of the Poynting vector components

In this section, potentiality of experimental analysis for the time-averaged Poynting vector is considered. In part, we will show that combined application of conventional interferometry and Stokes-polarimetry should allow of unambiguous determining characteristics of the time-averaged Poynting vector components at each point of electromagnetic field.



One of theoretical aspects in rapidly developing area of the modern optical technology, elaboration of new kinds of optical tweezers [25 - 30, 85] is connected with the fact that the vortex beams and polarized waves (both homogeneous and inhomogeneous) possess an angular momentum [85 - 87]. Existence of controlled angular momentum provides a possibility for controlled rotation of micro-objects locked by corresponding optical traps. Angular momentum of a field can be specified at each spatial point. One may also consider angular momentum averaged over some spatial area. As it is known [86, 88], angular momentum may be divided into a spin momentum associated with circular polarization, and an orbital one produced by specific beam structure. However, density of the angular momentum,  $j_z$ , (at least, of its orbital part) depends upon the location of the axis  $\mathbf{r}=0$  (i.e., "purchase") with regard to which the parameter  $j_z$  is calculated. As a result, some ambiguity appears. At the same time, another physical value closely associated with the angular momentum, *viz.* the space distribution of characteristics of the time-averaged Poynting vector (more accurately, its transverse component), represents a univocal function of coordinates of each field point.

Distribution of parameters of the time-averaged Poynting vector for Laguerre-Gaussian beams has been considered in [89, 90]. However, behavior of the time-averaged Poynting vector was analyzed only for homogeneously polarized fields and "symmetrical" beams [86, 89].

At the same time, analyzed fields may be more complicated, in part, when their polarization is inhomogeneous. Distribution of the transverse component of the Poynting vector for such fields may be characterized by a set of certain points, i.e., by the net of Poynting vector singularities [87]. The importance of these points comes from the fact that the characteristics of this singular net, such as the Poynting field skeleton, determine qualitative behavior of the Poynting vector at each of the field points [37]. For instance, vortex Poynting singularities [91] are the points, around which circulation of the transverse component of the Poynting vector takes place. In other words, one deals here with the points which are the intersections of the observation plane and the axis of the angular momentum. Thus, spatial distributions of characteristics of the time-averaged Poynting vector components would contain important information on the field, which is concerned with the energy flows [24, 30, 35, 36].

Nevertheless, one can state that no the technique for experimental analysis of the time-averaged Poynting vector components and their singularities has been developed up to now. Here we would like to demonstrate that the components of the time-averaged Poynting vector can be experimentally analyzed by using conventional optical methods.

It has been shown [37, 87] that the instantaneous components of the Poynting vector may be written as

$$\begin{cases} P_x \approx \frac{c}{4\pi k} \{E_x T_2 - E_y T_1\}; \\ P_y \approx \frac{c}{4\pi k} \{E_y T_2 + E_x T_1\}; \\ P_z \approx \frac{c}{4\pi} \{E_x^2 + E_y^2\}, \end{cases} \quad (14)$$

where

$$\begin{cases} T_1 = E_x \Phi_x^y - E_y \Phi_y^x + \frac{A_x^y}{A_x} E_{x, \frac{\pi}{2}} - \frac{A_y^x}{A_y} E_{y, \frac{\pi}{2}}; \\ T_2 = E_x \Phi_x^x + E_y \Phi_y^y + \frac{A_x^x}{A_x} E_{x, \frac{\pi}{2}} + \frac{A_y^y}{A_y} E_{y, \frac{\pi}{2}}, \end{cases} \quad (15)$$

$$\begin{cases} E_i = A_i \cos(\omega t + \Phi_i - kz); \\ E_{i, \frac{\pi}{2}} = A_i \sin(\omega t + \Phi_i - kz), \end{cases} \quad (16)$$

$i = x, y$ ,  $A_i, \Phi_i$  - denote, respectively, amplitudes and phases of the corresponding field components,  $A_i^l, \Phi_i^l$  are their partial derivatives, and  $k = \omega / c$  represents the wave number. Note, here the axis  $z$  coincides with the preferential direction of the wave propagation.

One can show after some algebraic transformations that the averaged components of the Poynting vector are as follows:

$$\begin{cases} \bar{P}_x \approx \frac{c}{8\pi k} \{ [A_x^2 \Phi_x^x + A_y^2 \Phi_y^x] - A_x A_y (\Phi_x^y - \Phi_y^x) \cos \Delta - (A_x A_y^y + A_y A_x^x) \sin \Delta \}; \\ \bar{P}_y \approx \frac{c}{8\pi k} \{ [A_x^2 \Phi_x^y + A_y^2 \Phi_y^y] + A_x A_y (\Phi_x^x - \Phi_y^x) \cos \Delta + (A_x A_x^x + A_y A_y^x) \sin \Delta \}; \\ \bar{P}_z \approx \frac{c}{8\pi} \{ A_x^2 + A_y^2 \}. \end{cases} \quad (17)$$

The second and the third terms in the brackets appearing of the first two rows of Eq. (17) may be rewritten as follows:

$$A_x A_y (\Phi_x^i - \Phi_y^i) \cos \Delta + (A_x A_y^i + A_y A_x^i) \sin \Delta = \frac{\partial}{\partial i} (A_x A_y \sin \Delta), \quad (18)$$

where  $i = x, y$ . Then, the system of Eqs. (17) is transformed to

$$\begin{cases} \bar{P}_x \approx -\frac{c}{8\pi k} \{ [A_x^2 \Phi_x^x + A_y^2 \Phi_y^x] - \frac{\partial}{\partial y} (A_x A_y \sin \Delta) \} \\ \bar{P}_y \approx -\frac{c}{8\pi k} \{ [A_x^2 \Phi_x^y + A_y^2 \Phi_y^y] + \frac{\partial}{\partial x} (A_x A_y \sin \Delta) \} \\ \bar{P}_z \approx \frac{c}{8\pi} \{ A_x^2 + A_y^2 \} \end{cases} \quad (19)$$

As it follows from Eqs. 19, one of the possible ways of the Poynting vector components measurement is complete analysis of orthogonal components:

Measurement of the components intensities (for the defining of components amplitudes and their corresponding derivatives).

Phasometry (interferometry) of the components (for the defining of components phases and their corresponding derivatives).

However, the phase difference between orthogonal components  $\Delta$  is included in relations (19). Note that:

- i. as it is known [17] measurement of the “absolute” phase in optics is “problematic”. Only phase difference between object wave and referent one is fixed. In other words, the phase is measured within the constant component.
- ii. the experimental arrangement for simultaneous measuring of component phase with the same constant components is practically impossible.

Consequently, the phase difference  $\Delta$  cannot be measured within the enough accuracy by the components analysis. Such value may be easily measured by the Stokes polarimetry [83]. Stocks parameters for monochromatic wave have the form [17]:

$$\begin{aligned} s_0 &= A_x^2 + A_y^2 \\ s_1 &= A_x^2 - A_y^2 \\ s_2 &= 2A_x A_y \cos \Delta \\ s_3 &= 2A_x A_y \sin \Delta. \end{aligned} \quad (20)$$

Value  $\Delta$  may be determined as:

$$\Delta = \tan^{-1} \frac{s_3}{s_2}, \quad (21)$$

and the Eqs. 19 may be used for the analysis of Poynting vector components. If one takes into account Eqs. (20), the relations (19) are transformed to the form:

$$\left\{ \begin{aligned} \bar{P}_x &\approx -\frac{c}{16\pi k} \{[(s_0 + s_1)\Phi_x^x + (s_0 - s_1)\Phi_y^x] - \frac{\partial s_3}{\partial y}\} \\ \bar{P}_y &\approx -\frac{c}{16\pi k} \{[(s_0 + s_1)\Phi_x^y + (s_0 - s_1)\Phi_y^y] + \frac{\partial s_3}{\partial x}\} \\ \bar{P}_z &\approx \frac{c}{8\pi} s_0 \end{aligned} \right. \quad (22)$$

Thus, components of the Poynting vector may be defined by the Stokes parameters and derivatives of components phases. Note, that phasometry of only one component (let us  $y$ -component for certainty) is necessary, because a phase of the orthogonal component is determined as  $\Phi_x = \Delta + \Phi_y$ . Obviously, in practical sense, the phasometry of the smooth component (without any singularity in the analyzed area) is preferable.

As it is known (see, for example, [92] which is the closest for the topic of our study), three kinds of measurement are necessary for defining of phase of scalar field at each field point:

Intensity of component field  $I_i = A_i^2$ .

Intensity of referent wave  $I_r = A_r^2$ .

Intensity of total field  $I_s$ .

The component phase (within the constant component) may be derived from the relation:

$$\Phi_i = \arccos \frac{I_s - (I_i + I_r)}{2\sqrt{I_i I_s}}. \quad (23)$$

Naturally, the phase derivatives  $\Phi_i^k$ , included in Eqs. (19), (22) are independent on the constant component.

Thus, joint applying of the conventional interferometry and Stokes-polarimetry allow us unambiguously determine the characteristics of the Poynting vector components at each point of an optical field.

## **6. Complex degree of mutual polarization of laser fields formed by birefringent matrices of biological tissues. Polarimetric metrology in diagnostics of structure and physiological state of biological tissues**

Optical diagnostics of complex, inhomogeneous in phase objects, such as biological tissue (BT) includes relatively independent directions, *viz.* photometric, polarimetric and correlation. On the base of these directions, *laser polarimetry of BTs* has been formed as a new and intensively developed branch of applied optics metrology [93 - 97]. Within this branch, one considers morphological structure of BTs as two-component amorphous-crystalline structure. Crystalline component or extracellular matrix of BTs of the main types is formed by the net of coaxial cylindrical protein fibrils possessing the properties of uniaxial birerigent crystals.

Laser polarimetry has been shown to be efficient tool under single-scattering regime for finding out the set of interconnections among statistical (the first to the fourth statistical moments) and fractal (log-log dependence of power spectra) parameters, which characterize coordinate geometrical-optical structure of extracellular matrix of BT, on the one hand, and polarization parameters (coordinate distributions of the azimuth of polarization and ellipticity) at its image, on the other hand. The mentioned success in diagnostic application of laser polarimetry of optically thin BT layers stipulates further development of new statistical approaches for analysis of concordance of the states of polarization at various points of an object field formed by multy-layer BT, in part by applying the concepts of the complex degree of mutual polarization (CDMP) [6, 32, 98, 99] and determining on this base the complex degree of mutual anisotropy [100] of polycrystalline protein nets. It is of vital importance for development of new techniques for reliable estimating physiological state of optically anisotropic BTs.

Spatial modulation of correlation and polarization characteristics is intrinsic to laser radiation scattered by object fields, including images of BTs [4, 5, 93, 101 - 103]. Optical coherent tomography [104, 105] and its new branch, polarization-sensitive optical coherent tomography [95, 96, 101, 106, 107], are among the most promising techniques for obtaining information on such modulation. This diagnostic technique is based on measuring coordinate distributions of the Stokes parameters at BT images that give a possibility to obtain important data both on BT's microstructure and on magnitude and coordinate distribution of optical anisotropy of architectonic nets formed by birefringent protein fibrils. Further progress of polarization-sensitive optical coherent tomography presumes development of new techniques for polarimetric, interferometric and correlation analysis and processing of inhomogeneous in polarization images of BTs. Let us consider these techniques in more details.

## 6.1 Polarimetric approach

Laser radiation, similarly to natural light, can be absorbed and scattered by BT. Each of these processes leads to enrichment of the field by information on micro- and macrostructure of the studied medium and its components. Spectrophotometric techniques are among the most widely used now for diagnostics of BTs. These techniques are based on the analysis of spatial and temporal changes of intensity of a field scattered by optically inhomogeneous samples. At the same time, other diagnostic techniques are also intensively developed, being based on fundamental concepts of *polarization* and *coherence*.

Spatial fluctuations of the parameters of optical fields are traditionally characterized in terms of the field's coherent properties [1]. The concept of *the measure of coherence* between two light disturbances is associated with ability of such disturbances to produce interference pattern and, consequently, with visibility of an interference pattern [17, 108]. It is just the measure of the sum of correlations between equal polarization projections of electrical fields at two specified points.

Another type of correlation characteristics of scattered laser fields is the degree of polarization defined as the maximal magnitude of correlations between the orthogonal polarization projections of a field at the fixed point [17]. Polarization properties of a field are experimentally investigated by measuring intensity of radiation passed various optically active elements of a medium. The techniques based on the use of the coherency matrix and the degree of polarization related to correlation of orthogonal in polarization components at one point of a field are often referred to as the polarization techniques [93].

Intense development of vector (polarization) approach in the study of morphological structure and physiological state of various BTs [96] formed the basis for elaborating model concepts on optically anisotropic and self-similar structure of BTs [98, 109 - 112]. So, Cowin [113] analyzes hierarchical self-similar structure of typical connective tissue, *viz.* tendon (tropocollagen, microfibril, subfibril, fascia etc.) Cowin emphasizes that threadlike structural elements are discrete, being characterized by scale recurrence over large interval of "optical sizes" (from 1  $\mu\text{m}$  to  $10^3 \mu\text{m}$ ). For that, optical characteristics of such structure of BT correspond, generally, to "frozen" optically uniaxial liquid crystals.

The same approach for describing morphological structure of BT has been applied in Refs [111, 112], where a BT is considered as two-component amorphous-crystalline structure. Amorphous component of BT (fat, lipids, unstructured proteins) is isotropic in polarization, i.e. optically inactive.

Crystalline component of BT is formed by oriented birefringent protein fibrils (collagen proteins, myosin, elastin etc.). The properties of the each isolated fibril are modeled by optically uniaxial crystal, whose axis direction coincides with the direction of packing at the plane of BT, and the birefringence coefficient is determined by the fibril matter. Architectonic net formed by disordered birefringent fibrils constitutes higher level of BT organization.

This model provides explanation of polarization inhomogeneity of object fields produced by BTs of various types, such as osseous and muscular tissues, tissues of female reproductive organs (myometrium) [111, 112]. The interconnections between the azimuth of polarization and ellipticity of a field, on the one hand, and the directions of packing of fibrils, as well as

the parameters characterizing birefringence, on the other hand, have been also determined for the single-scattering regime [113]. It allowed improving the technique of polarization visualizing the BT's architectonics by applying statistical analysis of 2D distributions of scattered fields [111, 112].

The papers [111, 114, 115] represent the results on determining the interconnections among the set of the 1<sup>st</sup> to the 4<sup>th</sup> order statistical moments characterizing microgeometry of a surface and orientation/phase structure of human BT's birefringent architectonics, on the one hand, and the set of the corresponding statistical moments of 2D distributions of the azimuth of polarization and ellipticity of the images of these objects, i.e. polarization maps, on the other hand. It has been stated that increasing asymmetry and excess characterizing the distributions of the azimuth of polarization and ellipticity at polarization maps result from increasing dispersion in orientation of birefringent fibril optical axes. Decreasing asymmetry and excess correspond to increasing dispersion of phase delays caused by biological crystals of architectonic nets.

Further development of laser polarimetry led to new techniques for measuring 2D arrays of polarization parameters that characterize nets of biological crystals inherent in various types of human's BTs. So, statistical analysis of the coordinate distributions of the Stokes parameters provide new information on microstructure (such as magnitudes and coordinate distributions of the parameters of optical anisotropy of architectonic nets formed by collagen or myosin) of physiologically normal and pathologically changed BTs [116]. Generally, intensive development of the techniques for diagnostic applications of laser radiation has been reflected into optical coherent tomography that became the most elaborated and convenient instrument for non-invasive study of BT structure.

The use of polarization of laser beam as a tool for contrasting of BT images resulted in a new branch of optical coherent tomography (OCT) [94, 104, 105], *viz.* polarization-sensitive optical coherent tomography (PSOCT) [95, 96, 101, 106, 107]. Note, special feature of *laser polarimetry of distributions of the azimuth of polarization and ellipticity* consists in point-by-point analysis of the object field's polarization parameters followed by searching for interconnections of these parameters with orientation and anisotropic parameters of BT's architectonics. For that, it leaves undetermined the data on peculiarities (statistical, fractal) of 2D distributions of polarization parameters of a field and orientation/phase characteristics of an object. So, further development of the techniques for non-invasive macro-diagnostics of geometrical optical structure of BT through improving conventional polarization-interference mapping and looking for new techniques for reconstruction of BT architectonics is among vital topics of modern optics.

## 6.2 Optical correlation approach

It is known [5, 32, 98, 108] that polarization properties of light at specified point of space can be described by the coherency matrix. This formalism is comprehensive for of a light field as a whole, when the field is statistically homogeneous, i.e. when the field's characteristics are independent on spatial coordinates. However, for spatially inhomogeneous fields it is of importance to know not only coordinate distributions of polarization parameters, but also interconnections of the states of polarization and the degree of coherence at different points of a field.

The first attempt to describe spatially inhomogeneous in polarization optical fields consisted in direct generalization of the coherency matrix to the two-point coherency (polarization) matrix performed by Gori [98, 117]. (Note, this principle of representation of optical fields of general type, i.e. partially spatially coherent and inhomogeneously polarized fields, even without restrictions connected with the paraxial approximation, has been explicitly formulated yet at the morning of the era of lasers (*'litic age'*), in the early sixties, being summed up in seminal review by Wolf and Mandel [1].) In *'post-litic age'*, Gori shows that some magnitude of interference pattern's visibility corresponds to the each coherency matrix element; these patterns result from superposition of radiations from two point sources whose polarization characteristics are formed by the set of polarizers and phase plates. Matrix analysis of correlation properties of scattered coherent radiation has been generalized for vector (inhomogeneously polarized) fields [109, 110].

As much prospective is the development of the tools for direct measuring CDMP in problems of biomedical optics connected with processing of coherent, inhomogeneous in polarization images of BTs obtained by allying the OCT techniques. It has been shown [100] that the CDMP of BT's coherent image is the parameter sensitive to orientation/phase changes of BT's architectonics. Experimental study [118] of 2D distributions of the CDMP of BT's laser images for examples of muscular, skin and osseous tissues have corroborated existence of the interconnections between the coordinate structure of the CDMP at laser images and geometrical/optical structure of birefringent architectonic nets of physiologically normal and dystrophically changed BTs. Taking into account diagnostic feasibilities of the CDMP, further searching for peculiarities of the coordinate distributions of this parameter for various BTs seems to be quite relevant and urgent problem. The following consideration is devoted to this problem.

### 6.3 BT as birefringent extracellular matrix transforming laser light parameters

As it has been mentioned above, BT consists of two components, *viz.* optically isotropic (amorphous) and anisotropic net (extracellular matrix) of birefringent optically uniaxial fibril [4, 94, 103, 111, 112], see Figure 20.

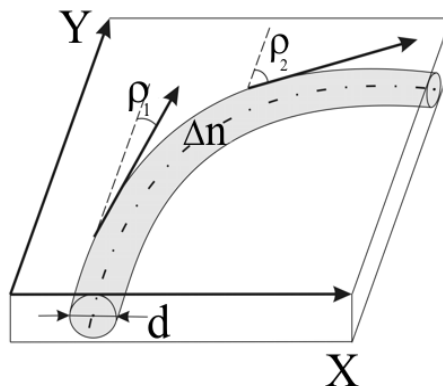


Fig. 20.  $\Delta n$  -  $d$ -diam birefringent fibril;  $\rho_i$  - the directions of the fibril packing at the plane of a BT sample.

The action of amorphous and architectonic (crystalline) components of BT,  $\{A\}$  and  $\{C\}$ , respectively, on coherent radiation is characterized by the following Jones matrix operators:

$$\{A\} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} \exp(-\tau l) & 0 \\ 0 & \exp(-\tau l) \end{vmatrix}; \quad (24)$$

$$\{C\} = \begin{vmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{vmatrix} = \begin{vmatrix} \cos^2 \rho + \sin^2 \rho \exp(-i\delta) & \cos \rho \sin \rho [1 - \exp(-i\delta)] \\ \cos \rho \sin \rho [1 - \exp(-i\delta)] & \sin^2 \rho + \cos^2 \rho \exp(-i\delta) \end{vmatrix}; \quad (25)$$

Here,  $\tau$  is the absorption coefficient for BT of thickness  $l$ ;  $\rho$  is the packing direction of anisotropic fibril (with birefringence coefficient  $\Delta n$ ) at the plane of a BT's sample introducing a phase shift  $\delta = 2\pi/\lambda \Delta n l$  between the orthogonal polarization components,  $E_x, E_y$ , of the probing laser beam of wavelength  $\lambda$ .

#### 6.4 Mechanisms of forming inhomogeneous in polarization BTs laser images

As it follows from analysis performed in Refs [93, 95, 97, 113] the mechanism of forming inhomogeneous in polarization boundary field of a BT at each point can be represented in the following form:

"decomposition" of an amplitude of laser wave  $U$  into orthogonal linearly polarized mutually coherent components:

$$\begin{pmatrix} U_x(r) \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ U_y(r) \end{pmatrix};$$

forming a phase shift (phase difference) between these components accounting birefringence,  $\delta(r)$ ;

superposing the orthogonally polarized components that results, in general case, in elliptically polarized wave that is described by the following equation:

$$\frac{X^2}{U_x^2(r)} + \frac{Y^2}{U_y^2(r)} - \frac{2XY}{U_x(r)U_y(r)} \cos \delta(r) = \sin^2 \delta(r). \quad (26)$$

#### 6.5 Statistical and fractal analysis of polarization images of BTs

Two types of optically thin (extinction coefficient  $\tau \leq 0.1$ ) of BT's histological tomes have been studied in Refs [4, 97]:

- structured osseous tissue (Fig. 21 a, b);
- parenchymatous tissue of kidney (Fig. 21 c, d).

Coordinate distributions  $\{\alpha(r); \beta(r)\}$  and histograms  $W(\alpha), W(\beta)$  of the magnitudes of the azimuth of polarization and ellipticity at images of histological tomes of physiologically normal osseous tissue (left part) and kidney tissue (right part) are shown in Figure 22 (fragments (a), (b) and (c), (d), respectively).



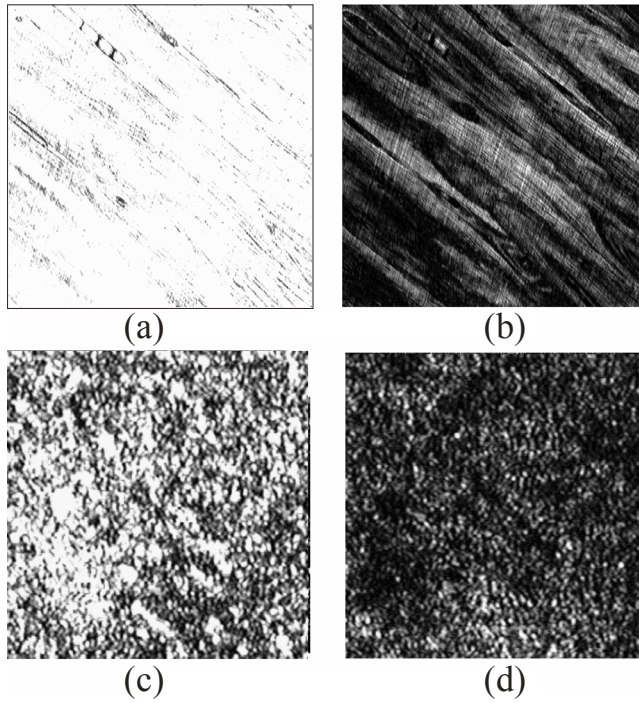


Fig. 21. Polarization images of osseous tissue (a, b) and tissue of kidney (c, d) for matched (a, c) and crossed (b, d) polarizer and analyzer, respectively.

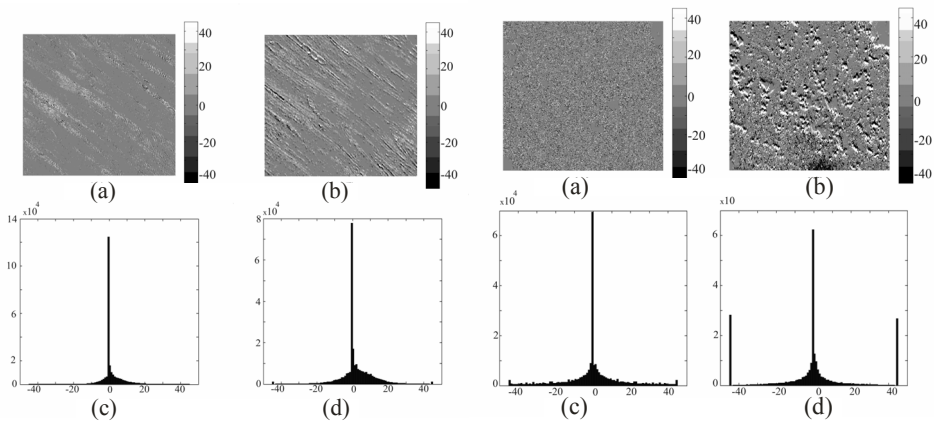


Fig. 22. Polarization maps of osseous tissue (left part) and kidney tissue (right part). Fragments (a), (b) correspond to 2D distributions of the azimuth of polarization and ellipticity, respectively; fragments (c), (d) show histograms of the corresponding distributions.

Distributions of the azimuth of polarization and ellipticity of the maps of BTs of two types characterize the set of statistical moments of the 1<sup>st</sup> to the 4<sup>th</sup> orders shown in Table 1.

The obtained data for statistical moments of the 1<sup>st</sup> to the 4<sup>th</sup> orders for distributions  $W(\alpha), W(\beta)$  of images of BTs of different morphological construction show that as birefringent architectonic nets are higher structured, as the magnitudes of the 3<sup>rd</sup> and the 4<sup>th</sup> statistical moments associated with the set of polarization parameters increase.

<i>Osseous tissue (31 samples)</i>				<i>Kidney tissue (27 samples)</i>			
	$\alpha(r)$		$\beta(r)$		$\alpha(r)$		$\beta(r)$
$M_1$	$0,38 \pm 0,027$	$M_1$	$0,24 \pm 0,014$	$M_1$	$0,11 \pm 0,01$	$M_1$	$0,08 \pm 0,004$
$M_2$	$0,25 \pm 0,015$	$M_2$	$0,21 \pm 0,017$	$M_2$	$0,19 \pm 0,013$	$M_2$	$0,05 \pm 0,003$
$M_3$	$9,8 \pm 0,882$	$M_3$	$7,7 \pm 0,539$	$M_3$	$1,4 \pm 0,056$	$M_3$	$0,61 \pm 0,04$
$M_4$	$24,6 \pm 2,71$	$M_4$	$12,5 \pm 1,125$	$M_4$	$3,1 \pm 0,093$	$M_4$	$2,25 \pm 0,113$

Table 1. Statistical momentums  $M_i$  of the coordinate distributions of the states of polarization at images of BTs of osseous and kidney BTs

## 6.6 Diagnostic feasibilities of polarization mapping of optically thin layers of BTs

To look for the feasibilities for differentiation of geometrical/optical structure of architectonics of BTs, comparative study of statistical and fractal structures of BT maps has been carried out [103, 119 - 121]. There are BTs of interest: (i) physiologically normal and dystrophically changed muscular tissue (MT), i.e. structured BT with ordered architectonics; (ii) physiologically normal and septically inflamed pulmonary tissue (PT), i.e. BT with island architectonics. The set of figures (Figs. 23 and 24) illustrates polarization maps of the mentioned objects.

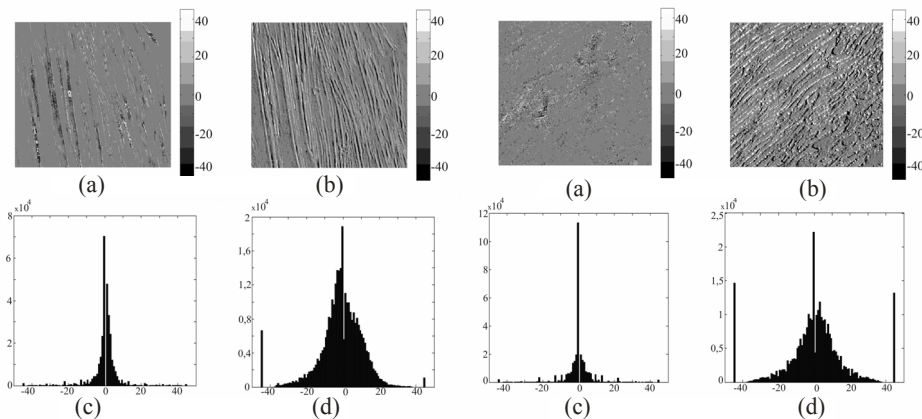


Fig. 23. Polarization maps of physiologically normal (left part) and dystrophically changed (right part) MT: (a), (b) - 2D distributions (in degrees) of the azimuth of polarization and ellipticity; (c), (d) - histograms of polarization parameters.

Potentialities for statistical differentiation of polarization maps are illustrated by comparison of magnitudes of the statistical moments of the 1<sup>st</sup> to the 4<sup>th</sup> orders for the azimuth of polarization and ellipticity at images of physiologically normal and pathologically changed BTs, represented in Tables 2 and 3.

<i>Norm (21 samples)</i>			<i>Pathology (22 samples)</i>		
$M_1$	$\alpha$	$\beta$	$M_1$	$\alpha$	$\beta$
$M_1$	$0,26 \pm 0,013$	$0,12 \pm 0,01$	$M_1$	$0,21 \pm 0,001$	$0,18 \pm 0,011$
$M_2$	$0,12 \pm 0,01$	$0,08 \pm 0,004$	$M_1$	$0,19 \pm 0,013$	$0,12 \pm 0,006$
$M_3$	$6,7 \pm 0,469$	$4,9 \pm 0,294$	$M_1$	$9,4 \pm 0,846$	$6,2 \pm 0,434$
$M_4$	$17,9 \pm 1,61$	$14,5 \pm 1,595$	$M_1$	$25,4 \pm 2,54$	$21,6 \pm 2,592$

Table 2. Statistical moments of the 1st to the 4<sup>th</sup> orders for the coordinate distributions of the azimuth of polarization and ellipticity at images of histological tomes of normal and dystrophically changed MT

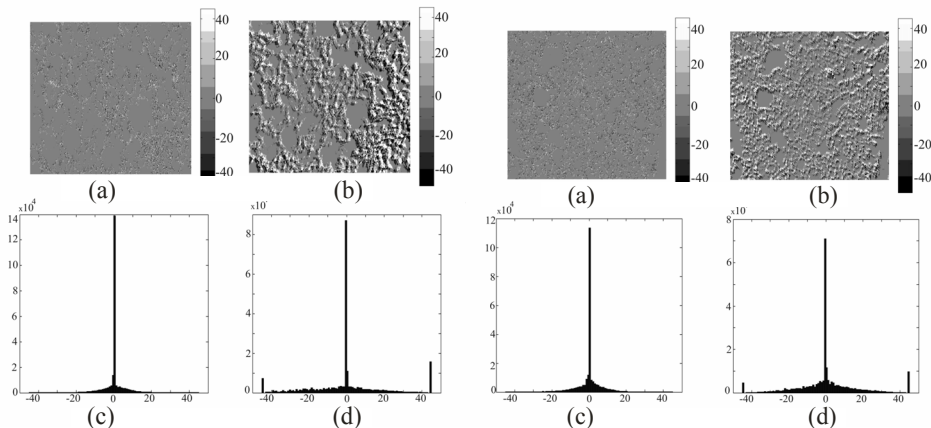


Fig. 24. Polarization maps of physiologically normal (left part) and septicly inflamed (right part) PT: (a), (b) - 2D distributions (in degrees) of the azimuth of polarization and ellipticity; (c), (d) - histograms of polarizarion parameters.

One can see from the represented data that magnitudes of asymmetry and excess of the coordinate distributions of the azimuth of polarization and ellipticity at BT images with ordered architectonics (Table 2) exceed by 3 to 5 times the magnitudes of the corresponding parameters characterizing statistics of the 3<sup>rd</sup> and the 4<sup>th</sup> orders of inhomogeneous in polarization BT images with island architectonics (Table 3).

Norm (22 samples)			Pathology (22 samples)		
$M_1$	$\alpha$	$\beta$	$M_1$	$\alpha$	$\beta$
$M_1$	$0,08 \pm 0,004$	$0,06 \pm 0,004$	$M_1$	$0,09 \pm 0,004$	$0,08 \pm 0,005$
$M_2$	$0,19 \pm 0,013$	$0,16 \pm 0,02$	$M_2$	$0,21 \pm 0,01$	$0,2 \pm 0,016$
$M_3$	$3,12 \pm 0,25$	$2,64 \pm 0,29$	$M_3$	$2,04 \pm 0,184$	$1,18 \pm 0,142$
$M_4$	$6,92 \pm 0,76$	$2,17 \pm 0,282$	$M_4$	$4,17 \pm 0,5$	$7,28 \pm 1,092$

Table 3. Statistical moments of the 1<sup>st</sup> to the 4<sup>th</sup> orders for the coordinate distributions of the azimuth of polarization and ellipticity at images of histological tomes of normal and septically inflamed PT

On the other hand, structured BTs possess hierarchical, self-similar structure of architectonics [97, 113, 120]. That is why one must determine optical manifestations of such geometry of anisotropic component of structured BTs. The study of self-similarity of polarization maps of BTs has been carried out for skin derma.

Figure 25 shows the set of log-log dependences,  $\log P(\alpha) - \log(1/d)$ ,  $\log P(\beta) - \log(1/d)$ , for power spectra of the distributions of polarization parameters at images of physiologically normal and pathologically changed skin derma.

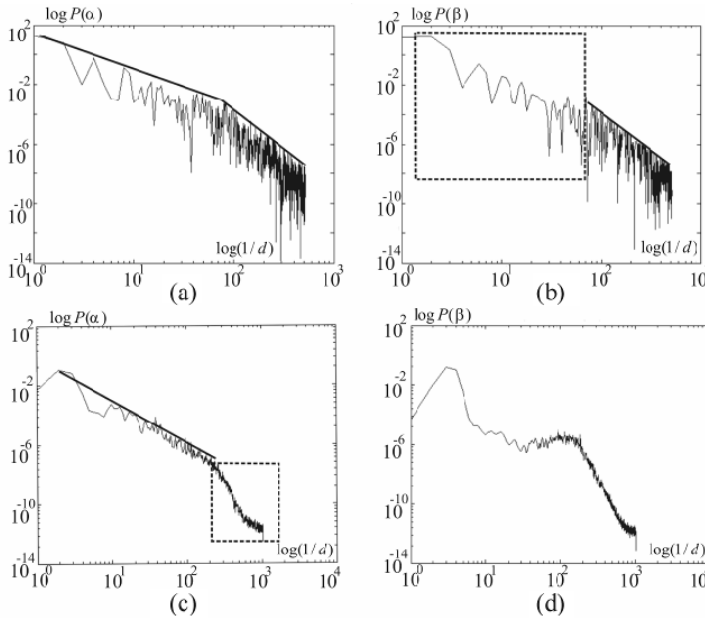


Fig. 25. log-log dependences of power spectra of the azimuth of polarization and ellipticity at images of physiologically normal (a), (b) and pathologically changed (c), (d) skin derma.

The dependence  $\log P(\alpha) - \log(1/d)$  manifests two slopes within three decades of scales of architectonics of skin derma's elements; the dependence  $\log P(\beta) - \log(1/d)$  consists of the part with constant slope and statistical part, as it is seen from Fig. 1.6 (a) and (b). Pathological state of such BT manifests itself in randomization of distributions of polarization parameters at images of the corresponding histological tomes, i.e. for approximating curves (Fig. 25 (c) and (d)) steady slope is absent.

### 6.7 Degree of mutual polarization of laser images of BTs

For laser light fields scattered by BTs, coordinate changes of polarization characteristics can be characterized by the complex degree of mutual polarization, CDMP, at two points  $(r_1; r_2)$  [6]:

$$W(r_1, r_2) = \frac{(U_x(r_1)U_x(r_2) - U_y(r_1)U_y(r_2)\exp(i(\delta_2(r_2) - \delta_1(r_1))))^2}{(U_x^2(r_1) + U_y^2(r_1))(U_x^2(r_2) + U_y^2(r_2))}. \quad (27)$$

The following interconnections between the real part of the CDMP,  $\text{Re}W \equiv \bar{W}(r_1, r_2)$ , and the magnitudes at the azimuth of polarization,  $\alpha(r)$ , ellipticity,  $\beta(r)$ , at points  $r$  of BT's image, and orthogonal components  $U_x = \sqrt{I_x}$ ,  $U_y = \sqrt{I_y}$  of the complex amplitude with phase difference between them  $\delta(r)$  have been established in study [100]:

$$\bar{W}(r_2, r_1) = \frac{\left( (I_x(r_2)I_x(r_1))^{\frac{1}{2}} - (I_y(r_2)I_y(r_1))^{\frac{1}{2}} \cos(\delta_2(r_2) - \delta_1(r_1)) \right)^2}{I(r_2)I(r_1)}, \quad (28)$$

where

$$\delta(r_1) = \tan^{-1} \left[ \frac{\tan[2\beta(r_1)]}{\tan[\alpha(r_1)]} \right]; \quad \delta(r_2) = \tan^{-1} \left[ \frac{\tan[2\beta(r_2)]}{\tan[\alpha(r_2)]} \right]. \quad (29)$$

### 6.8 Statistical approach to analysis of coordinate distributions of the CDMP at BT images

Normal and dystrophically changed tissues of skeleton muscle has been investigated in paper [100]. Coordinate distributions of the CDMP  $\bar{W}(x, y)$  at images of histological tomes of such samples are shown in Figure 26.

Statistical analysis of the coordinate distributions of magnitudes  $\bar{W}(x, y)$  shows for the map of tissue of skeleton muscle differences between the statistical moments of the 1<sup>st</sup> to the 3<sup>rd</sup> orders (within 30%-50%), while magnitudes of excess (the 4<sup>th</sup> order statistical moment) differ by 2 to 2.5 times.

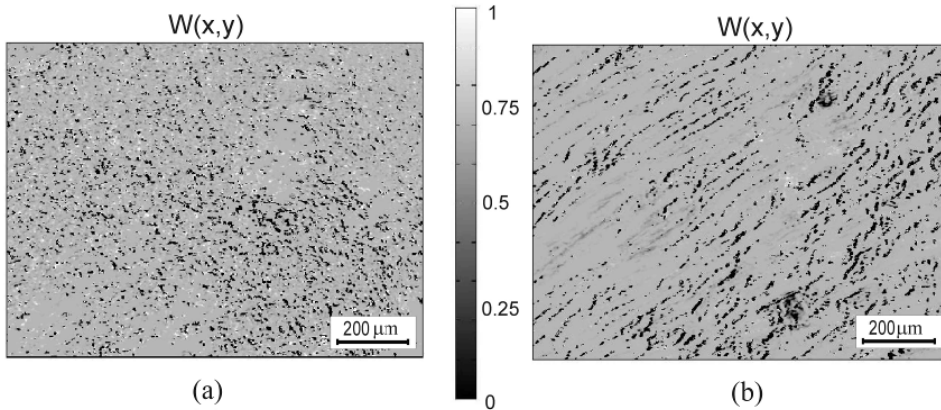


Fig. 26. Coordinate distributions of the CDMP at images of histological tomes of normal (a) and dystrophically changed (b) tissue of skeleton muscle.

Statistical moment	Muscular tissue (norm)	Muscular tissue (pathology)
	28 samples	25 samples
$M_1$	$0,62 \pm 0,03$	$0,49 \pm 0,02$
$M_2$	$0,1 \pm 0,06$	$0,14 \pm 0,01$
$M_3$	$32 \pm 3,2$	$21 \pm 2,31$
$M_4$	$81 \pm 8,72$	$43 \pm 3,87$

Table 4. Statistical moments of the 1<sup>st</sup> to the 4<sup>th</sup> orders for distributions of magnitudes  $\bar{W}(x,y)$  at images of histological tomes of normal and dystrophically changed tissues of skeleton muscle.

Figure 27 shows autocorrelation functions of the coordinate distributions of magnitudes of the CDMP at images of histological tomes of samples of normal (a) and pathologically changed (b) muscular tissues.

Parameter	Muscular tissue	
	Norm	Pathology
$L$	$0,25 \pm 0,02$	$0,13 \pm 0,01$
$\Omega$	$0,07 \pm 0,005$	$0,005 \pm 0,0001$

Table 5. Correlation parameters of the distributions of the CDMP,  $\bar{W}(x,y)$ , for normal and pathologically changed osseous muscular tissues.

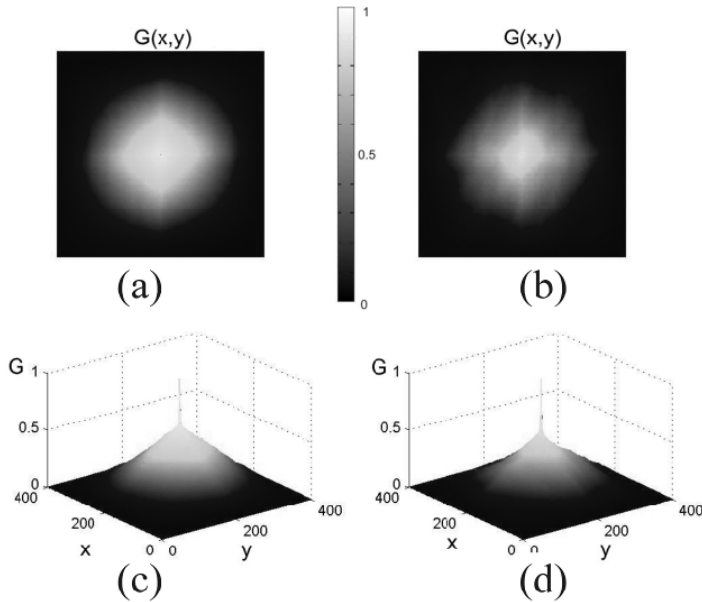


Fig. 27. Autocorrelation functions  $G(x,y)$  of 2D distributions of the CDMP,  $\bar{W}(x,y)$ , at images of histological tomes of muscular tissues.

Correlation analysis of 2D distributions of magnitudes  $\bar{W}(x,y)$  leads to the following conclusions: (i) differences between magnitudes of a half-width  $L$  of autocorrelation functions for distributions  $\bar{W}(x,y)$  at polarization/correlation maps of muscular tissue's maps are within the interval 0.65 to 2 times; (ii) differences between dispersions of the distributions for extrema of power spectra for  $\bar{W}(x,y)$  reach one order of magnitude.

Thus, the analysis of approaches and metrological techniques of modern polarimetric diagnostics of birefringent nets of BTs leads to the following conclusions.

- higher-order statistical moments (asymmetry and excess) characterizing the distributions of polarization/correlation parameters of laser images of BTs are of the highest sensitivity in respect to changing optical properties of nets of biological crystals of normal and pathologically changed BTs;
- interconnection between the real part of the CDMP and parameters of anisotropy (such as direction of an optical axis and phase difference between orthogonal components of scattered laser radiation) at different points of birefringent matrix provides reliable tool for correlation optical diagnostics of pathological changes of biological tissues and, as a consequence, for early and non-invasive diagnostics of wide-spread diseases, which challenge Humanity in the Third Millennium.

## 7. Conclusions

Thus, new approaches to metrology of partially coherent and partially polarized light fields are derived proceeding from singular optics concept. The first of them concerns to exploring

the spatial modulated time-averaged Poynting vector in completely and partially coherent non-paraxial light fields for control the motion of nano- and microparticles in optical currents. The second approach reveals interconnection between polarization singularities inherent in completely polarized and partially polarized optical beams for the general case of partially mutual coherence of orthogonally polarized components. Further, we have shown the feasibility for experimental analysis of the Poynting vector components by combining interferometric and polarimetric metrological techniques. At last, we have demonstrated potentialities of laser 2D Stokes-polarimetric metrological technique developed and implemented in previous consideration in non-invasive pre-clinic diagnostics of physiological state of biological tissues. The represented approaches show fruitfulness of attracting the concepts and metrological tools of singular optics in formation and investigation of unconventional polarization distributions that can be of usefulness in problems of optical correlation diagnostics and optical telecommunications.

There are the important next steps in developing the direction of Metrology represented in this chapter. The study of optical currents using light-scattering particles of nano-scale will provide vital data both on microstructure of light and on intimate processes of interaction of optical beams with extremely small (in scale of a wave length) particles and ensembles of them. It is in prospect to investigate experimentally in more details influence of varying degree of coherence of optical fields on some phenomena considered in this review and associated with peculiarities (*singularities*) of the Poynting vector in complex inhomogeneous optical fields beyond the paraxial approximation. Also, it is important to obtain experimental verification, *viz.* visualization, of the mechanical action of the spin moment in optical beams. The study of influence of polarization characteristics of light field on motion of nanoparticles will be continued, as well as development of biomedical application of notion and techniques of completely/partially coherent, inhomogeneously polarized fields.

Another branch of the following R&D must be focused, in our opinion, on overcoming some disadvantages intrinsic to digital optical data surveying due to automatic gain control applied in the most of register tools, including CCD cameras. Really, dealing with differentiation metrology (such as 2D Stokes polarimetry) one wants to have comparable data in the set of measurements that is not always achievable with self-controlled in sensitivity digital registering devices. As compromise, one can consider combination of analog data recording and digital processing of these data. Nevertheless, this important point is worthy separate investigation for providing higher accuracy and reliability of modern optical metrology.

## 8. Acknowledgement

This work was supported by the Ministry of Education and Science of Ukraine, Grants No 0110U000188 and No 0111U000719.

## 9. References

- [1] Wolf, E. & Mandel, L. (1965). Coherence properties of optical fields, *Rev. Mod. Phys.* V. 37, (1965), pp. 231-287



- [2] Soskin, M.S. & Vasnetsov, M.V. (2001). Singular Optics, In: *Progress in Optics*, Wolf, E., V. 42, (2001), pp. 219-276
- [3] Gbur, G. & Visser, T.D. (2010). The Structure of Partially Coherent Fields, In: *Progress in Optics*, Wolf, E., V. 55, Issue C, (2010), pp. 285-341
- [4] Angelsky, O.V.; Ushenko, A.G.; Ushenko, Yu.A. & Pishak, V.P. (2007). Statistical and Fractal Structure of Biological Tissue Mueller Matrix Images, In: *Optical Correlation Techniques and Applications*, Angelsky, O.V., pp. 213-266, Washington: Society of Photo-Optical Instrumentation Engineers
- [5] Wolf, E. (2003). Unified theory of coherence and polarization of random electromagnetic beams, *Phys. Lett. A*, V. 312, (2003), pp. 263-267
- [6] Ellis, J. & Dogariu, A. (2004). Complex degree of mutual polarization, *Opt. Lett.* V. 29, (2004), pp. 536-538
- [7] Mandel, L. (1998). Anticoherence, *Pure Appl. Opt.* V. 7, (1998), pp. 927-932
- [8] Collier, R.; Burckhardt, L. & Lin, L. (1971). *Optical Holography*, Academic, New York
- [9] Goodman, J.W. (1985). *Statistical Optics*, Wiley-Interscience, New York
- [10] Glauber, R.J. (1962). *Lectures in Theoretical Physics* (4), Brittin, W.E.; Downs, J. (Eds.) Interscience, New York
- [11] Glauber, R.J. (1963). The quantum theory of optical coherence, *Phys Rev.* V. 6, (1963), pp. 2529-2539
- [12] Glauber, R.J. (1965). Optical Coherence and Photon Statistics, In: *Quantum Optics and Electronics*, deWitt, C.; Blandin, A. & Cohen-Tannoudji, C., pp. 65-185, Gordon & Breach, New York
- [13] Dirac, P.A.M. (1985). *Principles of Quantum Mechanics*, (4), Oxford U. Press, Oxford
- [14] Wolf, E. (1987). Non-cosmological redshifts of spectral lines, *Nature*, V. 326, (1987), pp. 363-365
- [15] Wolf, E. & James, D.F.V. (1996). Correlation-induced spectral changes, *Rep. Progr. Phys.*, V. 59, (1996), pp. 771-818
- [16] Polyanskii, V.K. & Kovalskii, L.V. (1973). On fine structure of the scattered radiation field, *Opt. Spectrosc.*, V. 35, (1973), pp. 345-350
- [17] Born, M. & Wolf, E. (1999). *Principles of Optics*, (7), (expanded), Cambridge U. Press, Cambridge
- [18] Nye, J.F. (1999). *Natural Focusing and Fine Structure of Light: Caustics and Wave Dislocations* Institute of Physics Publishing, Bristol and Philadelphia
- [19] Freund, I. (2004). Coherency matrix description of optical polarization singularities, *J. Opt. A: Pure Appl. Opt.*, V. 6, (2004), pp. 229-234
- [20] Dolgov, A.D.; Doroshkevich, A.G.; Novikov, D.I. & Novikov, I.D. (1999). Classification of singular points in the polarization of the cosmic microwave background and eigenvectors of the Stokes matrix, *JETP Lett.*, V. 69, (1999), pp. 427-433
- [21] Hawking, S. (1988). *A Brief History of Time*, Bantam Press, ISBN 055305340X
- [22] Van de Hulst, H.C. (1957). *Light Scattering by Small Particles*, Wiley, New York
- [23] Bohren, C.F. & Huffman, D.R. (1983). *Absorbing and Scattering of Light by Small Particles*, Wiley, New York
- [24] Berry, M.V. (2009). Optical currents, *J. Opt. A: Pure Appl. Opt.* V. 11, (2009), 094001
- [25] He, H.; Friese, M.E.J.; Heckenberg, N.R. & Rubinsztein-Dunlop, H. (1995). Direct observation of transfer of angular momentum to absorptive particles from a laser beam with a phase singularity, *Phys. Rev. Lett.*, V. 75, (1995), pp. 826-829

- [26] Simpson, N.B.; Dholakia, K.; Allen, L. & Padgett, M.J. (1997). Mechanical equivalence of spin and orbital angular momentum of light: an optical spanner, *Opt. Lett.*, V. 22, (1997), pp. 52-54
- [27] O'Neil, A.T.; MacVicar, I.; Allen, L. & Padgett, M.J. (2002). Intrinsic and extrinsic nature of the orbital angular momentum of a light beam, *Phys. Rev. Lett.*, V. 88, (2002), 053601
- [28] Garces-Chavez, V.; McGloin, D.; Summers, M.D.; Fernandez-Nieves, A.; Spalding, G.C.; Cristobal, G. & Dholakia, K. (2004). The reconstruction of optical angular momentum after distortion in amplitude, phase and polarization, *J. Opt. A: Pure Appl. Opt.*, V. 6, (2004), pp. 235-238
- [29] Khonina, S.N.; Kotlyar, V.V.; Skidanov, R.V.; Soifer, V.A.; Jefimovs, K.; Simonen, J. & Turunen, J. (2004). Rotation of microparticles with Bessel beams generated by diffractive elements, *J. Mod. Opt.*, V. 51, (2004), pp. 2167-2184
- [30] Bekshaev, A.; Soskin, M. & Vasnetsov, M. (2008). *Paraxial Light Beams with Angular Momentum*, Nova Science Publishers, New York
- [31] Friberg, A.; Gao, C.; Eppich, B. & Weber, H. (1997). Generation of partially coherent fields with twist, *Proc. of SPIE 1997*, V. 3110, pp. 317-328
- [32] Tervo, J.; Setala, T. & Friberg, A. (2003). Degree of coherence for electromagnetic fields, *Opt. Express.*, V. 11, (2003), pp. 1137-1143
- [33] Setala, T.; Tervo, J. & Friberg, A.T. (2004). Complete electromagnetic coherence in the space-frequency domain, *Opt. Lett.*, V. 29, (2004), 328330
- [34] Tervo, J.; Setala, T. & Friberg, A.T. (2004). Theory of partially coherent electromagnetic fields in the space-frequency domain, *J. Opt. Soc. Am. A*, V. 21, (2004), pp. 2205-2215
- [35] Bekshaev, A.Ya. (2009). Oblique section of a paraxial light beam: criteria for azimuthal energy flow and orbital angular momentum, *J. Opt. A: Pure Appl. Opt.*, V. 11, (2009), 094003
- [36] Bekshaev, A.Ya.; Bliokh, K.Y. & Soskin, M.S. (2011). Internal flows and energy circulation in light beams, *J. Opt.*, V. 13 053001 (32pp) doi:10.1088/2040-8978/13/5/053001
- [37] Angelsky, O.V. (Ed). (2007). *Optical Correlation Techniques and Applications*, TA 1630, A6, SPIE Press, Bellingham, Washington
- [38] Angelsky, O.V.; Polyanskii, P.V. & Maksimyak, P.P. (2008). Speckles and Phase Singularities in Polychromatic Fields, In: *New Directions in Holography and Speckle*, Caulfield, H.J.; Vikram, Ch.S., pp. 37-53 (Chapter 3), Kluwer Academic Publishers, Boston
- [39] Angelsky, O.V.; Dominikov, N.N.; Maksimyak, P.P. & Tudor, T. (1999). Experimental revealing of polarization waves, *Appl. Opt.*, V. 38, (1999), pp. 3112-3117
- [40] Tudor, T. (1997). Polarization waves as observable phenomena, *J. Opt. Soc. Am. A*, V. 14, (1997), pp. 2013-2020
- [41] Berry, M.V. & Donald, K.T. (2008). Exact and geometrical optics energy trajectories in twisted beams, *J. Opt. A: Pure Appl. Opt.*, V. 10, (2008), 035005
- [42] Berry, M.V. & Dennis, M.R. (2001). Polarization singularities in isotropic random vector waves, *Proc. R. Soc., A*. 456, (2001), pp. 2059-2079

- [43] Angelsky, O.V.; Hanson, S.G.; Zenkova, C.Yu.; Gorsky, M.P. & Gorodys'ka, N.V. (2009). On polarization metrology (estimation) of the degree of coherence of optical waves, *Opt. Expr.*, V. 17, (2009), pp. 15623-15634
- [44] Zenkova, C.Yu.; Gorsky, M.P.; Maksimyak, P.P. & Maksimyak, A.P. (2010). Optical currents in vector fields, *Appl. Opt.*, V. 50, (2010), pp. 1105-1112
- [45] Mujait, M.; Dogariu, A. & Wolf, E. (2004). A law of interference of electromagnetic beams of any state of coherence and polarization and the Fresnel-Arago interference laws, *J. Opt. Soc. Am.*, A 21, (2004), pp. 2414-2417
- [46] Ellis, J.; Dogariu, A.; Ponomarenko, S. & Wolf, E. (2005). Degree of polarization of statistically stationary electromagnetic fields, *Opt. Commun.*, V. 248, (2005), pp. 333-337
- [47] Khrobatin, R. & Mokhun, I. (2008). Shift application point of angular momentum in the area of elementary polarization singularity, *J. Opt. A: Pure Appl. Opt.*, V. 10, (2008), 064015
- [48] Khrobatin, R.; Mokhun, I. & Victorovskaya, Ju. (2008). Potentiality of experimental analysis for characteristics of the Poynting vector components, *Ukr. J. Phys. Opt.*, V. 9, (2008), pp. 182-186
- [49] Bekshaev, A.Ya. & Soskin, M.S. (2007). Transverse energy flows in vectorial fields of paraxial light beams, *Proc. of SPIE 2007*, V. 6729, 67290G
- [50] Angelsky, O.V.; Yermolenko, S.B.; Zenkova, C.Yu. & Agelskaya, A.O. (2008). Polarization manifestations of correlation (intrinsic coherence) of optical fields, *Appl. Opt.*, V. 47, (2008), pp. 5492-5499
- [51] Angelsky, O.V.; Zenkova, C.Yu.; Gorsky, M.P. & Gorodys'ka, N.V. (2009). On the feasibility for estimating the degree of coherence of waves at near field, *Appl. Opt.*, V. 48, (2009), pp. 2784-2788
- [52] Angelsky, O.V.; Gorsky, M.P.; Maksimyak, P.P.; Maksimyak, A.P.; Hanson, S.G. & Zenkova, C.Yu. (2011). Investigation of optical currents in coherent and partially coherent vector fields, *Opt. Expr.*, V. 19, (2011), pp. 660-672
- [53] Turkevich, J.; Stevenson, P.C. & Hiller, J. (1951). A study of the nucleation and growth processes in the synthesis of colloidal gold, *J. Discuss. Faraday Soc.*, V. 11, (1951), pp. 55-75
- [54] Angelsky, O.V.; Hanson, S.G.; Maksimyak, P.P.; Maksimyak, A.P. & Negrych, A.L. (2008). Experimental demonstration of singular-optical colouring of regularly scattered white light, *J. Europ. Opt. Soc. Rap. Public.*, V. 3, (2008), 08029
- [55] Angelsky, O.V.; Hanson, S.G.; Maksimyak, A.P. & Maksimyak, P.P. (2005). Feasibilities of interferometric and chromoscopic techniques in study of phase singularities, *Appl. Opt.*, V. 44, (2005), pp. 5091-5100
- [56] Angelsky, O.V.; Maksimyak, A.P.; Maksimyak, P.P. & Hanson, S.G. (2005). Interference diagnostics of white-light vortices, *Opt. Expr.*, V. 13, (2005), pp. 8179-8183
- [57] Angelsky, O.V.; Hanson, S.G.; Maksimyak, A.P. & Maksimyak, P.P. (2005). On the feasibility for determining the amplitude zeroes in polychromatic fields, *Opt. Expr.*, V. 13, (2005), pp. 4396-4405
- [58] Angelsky, O.V.; Hanson, S.G.; Maksimyak, A.P. & Maksimyak, P.P. (2006). Optical correlation diagnostics of rough surfaces with large surface inhomogeneities, *Opt. Expr.*, V. 14, (2006), pp. 7299-7311

- [59] Angelsky, O.V.; Polyanskii, P.V. & Hanson, S.G. (2006). Singular-optical coloring of regularly scattered white light, *Opt. Expr.*, V. 14, (2006), pp. 7579-7586
- [60] Angelsky, O.V.; Bogatyryova, H.V. & Polyanskii, P.V. (2008). Interference coloring of regularly scattered white light, *Opt. Appl.*, V. 38, (2008), pp. 431-444
- [61] Felde, Ch.V.; Chernyshov, A.A.; Bogatyryova, H.V.; Polyanskii, P.V. & Soskin, M.S. (2008). Polarization singularities in partially coherent combined beams, *JETP Lett.*, V. 88, (2008), pp. 418-422
- [62] Chernyshov, A.A.; Felde, K.V.; Bogatyreva, G.V.; Polyanskii, P.V. & Soskin, M.S. (2009). Vector singularities of superposition of mutually incoherent orthogonally polarized beams, *Opt. Spectr.*, V. 107, (2009), pp. 645-650
- [63] Chernyshov, A.A.; Felde, Ch.V.; Bogatyryova, H.V.; Polyanskii, P.V. & Soskin, M.S. (2009). Vector singularities of the combined beams assembled from mutually incoherent orthogonally polarized components, *J. Opt. A: Pure Appl. Opt.*, V. 11, (2009), 094010
- [64] Polyanskii, P.V.; Felde, Ch.V. & Chernyshov, A.A. (2010). Polarization degree singularities, *Proc. of SPIE 2010*, V. 7388, 7388OA
- [65] Soskin, M.S. & Polyanskii, P.V. (2010). New polarization singularities of partially coherent light beams, *Proc. of SPIE 2010*, V. 7613, 7613OG
- [66] Polyanskii, P.V. & Bogatyryova, G.V. (2001). EDW: edge diffraction wave, edge dislocation wave, or whether *tertio est datur?*, *Proc. of SPIE*, V. 4607, pp. 109-124
- [67] Bogatyryova, G.V.; Felde, Ch.V.; Polyanskii, P.V.; Ponomarenko, S.A.; Soskin, M.S. & Wolf, E. (2003). Partially coherent vortex beams with a separable phase, *Opt. Lett.*, V. 28, (2003), pp. 878-880
- [68] Soskin, M.S.; Polyanskii, P.V., & Arkhelyuk, O.O. (2004). Computer-synthesized hologram-based rainbow optical vortices, *New J. Phys.*, V. 6, (2004), p. 196
- [69] Berry, M.V., & Dennis, M.R. (2004). Quantum cores of optical phase singularities, *J. Opt. A.: Pure Appl. Opt.*, V. 6, (2004), p.178
- [70] Freund, I.; Mokhun, A.I.; Soskin, M.S.; Angelsky, O.V. & Mokhun, I.I. (2002). Stokes singularity relations, *Opt. Lett.*, V. 27, (2002), pp. 545-547
- [71] Angelsky, O.V.; Mokhun, I.I.; Mokhun, A.I. & Soskin, M.S. (2002). Interferometric methods in diagnostics of polarization singularities, *Phys. Rev. E.*, V. 65, (2002), 036602(5)
- [72] Angelsky, O.V.; Mokhun, A.I.; Mokhun, I.I. & Soskin, M.S. (2002). The relationship between topological characteristics of component vortices and polarization singularities, *Opt. Commun.*, V. 207, (2002), pp. 57-65
- [73] Ponomarenko, S.A. (2001). A class of partially coherent vortex beams carrying optical vortices, *J. Opt. Soc. Am. A*, V. 18, (2001), 150-156
- [74] Gbur, G. & Visser, T.D. (2003). Coherence vortices in partially coherent beams, *Opt. Commun.*, V. 222, (2003), pp. 117-125
- [75] Schouten, H.F.; Gbur, G.; Visser, T.D. & Wolf, E. (2003). Phase singularities of the coherence functions in Young's interference pattern, *Opt. Lett.*, V. 28, (2003), pp.968-970
- [76] Swartzlander, G.A. Jr. & Schmit, J. (2004). Temporal correlation vortices and topological dispersion, *Phys. Rev. Lett.*, V. 93, (2004), 093901
- [77] Gbur, G.; Visser, T.D. & Wolf, E. (2002). Anomalous behavior of spectra near phase singularities of focused waves, *Phys. Rev. Lett.*, V. 88, (2002), 013901

- [78] Popescu, G. & Dogariu, A. (2002). Spectral anomalies at wave-front dislocations, *Phys. Rev. Lett.*, V. 88, (2002), 183902
- [79] Berry, M.V. (2002). Colored phase singularities, *New J. Phys.*, V. 4, (2002), p. 66
- [80] Berry, M.V. (2002). Exploring the colors of dark light, *New J. Phys.*, V. 4, (2002), p. 74
- [81] Volyar, A.V. & Fadeyeva, T.A. (2003). Generation of singular beams in uniaxial crystals, *Opt. Spectr.*, V. 94, (2003), pp. 260-270
- [82] Egorov, Yu.A.; Fadeyeva, T.A. & Volyar, A.V. (2004). The fine structure of singular beams in crystals: colors and polarization, *J. Opt. A.: Pure Appl. Opt.*, V. 6, (2004), p. 217
- [83] Azzam, R.M.A. & Bashara, N.M. (1977). *Ellipsometry and Polarized Light*, North-Holland, Amsterdam
- [84] Shurcliff, W.A. (1962). *Polarized Light: Production and Use*, Harvard Univ. Press, Cambridge, Massachusetts
- [85] Lang, M.J. & Block, S. M. (2003). Resource Letter: LBOT-1: Laser-based optical tweezers, *Am. J. Phys.*, V. 71, (2003), pp. 201-205
- [86] Allen, L.; Padgett, M.J. & Babiker, M. (1999). The orbital angular momentum of light, In: *Progress in Optics*, Wolf, E., V. 39, (1999), pp. 292-372
- [87] Mokhun, I.; Brandel, R. & Viktorovskaya, Ju. (2006). Angular momentum of electromagnetic field in areas of polarization singularities, *Ukr. J. Phys. Opt.*, V. 7, (2006), pp. 63-73
- [88] Berry, M.V. (1998). Paraxial beams of spinning light, *Proc. of SPIE 1998*, V.3487
- [89] Allen, L. & Padgett, M.J. (2000). The Poynting vector in Laguerre-Gaussian beams and the interpretation of their angular momentum density, *Opt. Comm.*, V. 198, (2000), pp. 67-71
- [90] Berry, M.V. (2005). Phase vortex spirals, *J. Phys. A*, V. 38, (2005), pp. 745-751
- [91] Mokhun, I.; Khrobatin, R.; Mokhun, A. & Viktorovskaya, Ju. (2007). The behavior of the Poynting vector in the area of elementary polarization singularities, *Opt. Appl.*, V. 37, (2007), pp. 261-277
- [92] Shvartsman, N. & Freund, I. (1995). Speckle spots ride phase saddles side saddle, *Opt. Com.*, V. 117, (1995), pp. 228-234
- [93] Jarry, G.; Steimer, E.; Damaschini, V.; Epifanie, M.; Jurczak, M. & Kaiser, R. (1998). Coherence and polarization of light propagating through scattering media and biological tissues, *Appl. Opt.*, V. 37, (1998), 7357
- [94] De Boer, J.F.; Milner, T.E.; Ducros, M.G.; Srinivas, S.M. & Nelson, J.S. (2002). Polarization-sensitive optical coherence tomography, In: *Handbook of Optical Coherence Tomography*. Bouma B.E.; Tearney G.J., pp. 237-274, Marcel Dekker Inc.: New York
- [95] Zimnyakov, D.A. & Tuchin V.V.(2002). Optical tomography of tissues, *Quantum Electronics*, V.32, (2002), pp. 849-867
- [96] Fercher, A. F. (2003). Optical coherence tomography - principles and applications, *Rep. Prog. Phys.*, V. 66, (2003), pp. 239-303.
- [97] Ushenko, A.G. & Pishak, V.P. (2004). Laser Polarimetry of Biological Tissue. Principles and Applications, In: *Coherent-Domain Optical Methods. Biomedical Diagnostics, Environmental and Material Science*, Tuchin. V., p. 67, Kluwer Academic Publishers
- [98] Gori, F.; Santarsiero, M.; Vicalvi, S.; Borghi, R. & Guattari, G. (1998). Beam coherence-polarization matrix, *Pure Appl. Opt.*, V. 7, (1998), pp. 941-951

- [99] Mujait, C. & Dogariu, A. (2004). Statistics of partially coherent beams: a numerical analysis, *J. Opt. Soc. Am. A.*, V 21, No. 6, (2004), pp. 1000-1003
- [100] Angelsky, O.V.; Ushenko, A.G. & Ushenko, Ye.G. (2005). Complex degree of mutual polarization of biological tissue coherent images for the diagnostics of their physiological state, *J Biomed Opt.*, V. 10. No. 6, (2005), pp. 060502
- [101] Everett, M.J.; Shoenenberger, K.; Colston, B.W. & Da Silva, L.B. (1998). Birefringence characterization of biological tissue by use of optical coherence tomography, *Opt. Lett.*, V. 23, (1998), pp. 228-230
- [102] Mujait, M. & Dogariu, A. (2003). Polarimetric and spectral changes in random electromagnetic fields, *Opt. Lett.*, V. 28, (2003), pp. 2153-2155
- [103] Angelsky, O.V.; Ushenko, A.G.; Ushenko, Yu.A.; Pishak, V.P. & Peresunko, A.P. (2010). Statistical, Correlation, and Topological Approaches in Diagnostics of the Structure and Physiological State of Birefringent Biological Tissues, In: *Handbook of Photonics for Biomedical Science*, Tuchin, V.V., pp. 21-67, USA: CRC Press
- [104] De Boer, J.F.; Milner, T.E.; Van Gemert, M.J.C. & Nelson, J.S. (1997). Two-dimensional birefringence imaging in biological tissue by polarization-sensitive optical coherence tomography, *Opt. Lett.*, V. 22, (1997), pp. 934-936
- [105] De Boer, J.F.; Milner, T.E. & Nelson, J.S. (1999). Determination of the depth-resolved Stokes parameters of light backscattered from turbid media by use of polarization-sensitive optical coherence tomography, *Opt. Lett.*, V. 24, (1999), pp. 300-302
- [106] Saxer, C.E.; de Boer, J.F.; Park, B.H.; Zhao, Y.; Chen, Z. & Nelson, J. S. (2000). High-speed fiber based polarization-sensitive optical coherence tomography of in vivo human skin, *Opt. Lett.*, V. 25, (2000), pp. 1355-1357
- [107] De Boer, J.F. & Milner, T.E. (2002). Review of polarization sensitive optical coherence tomography and Stokes vector determination, *J. Biomed. Opt.*, V. 7, (2002), pp. 359-371
- [108] Wolf, E. (1959). Coherence properties of partially polarized electromagnetic radiation, *Nuovo Cimento.*, V. 13, (1959), pp. 1165-1181
- [109] Mujait, M.; Dogariu, A. & Agarwal, G.S. (2004). Correlation matrix of a completely polarized, statistically stationary electromagnetic field, *Opt. Lett.*, V. 29, (2004), pp. 1539-1541
- [110] Korotkova, O. & Wolf, E. (2004). Spectral degree of coherence of a random three-dimensional electromagnetic field, *J. Opt. Soc. Am. A.*, V. 21, No.10, (2004), pp. 2382-2385
- [111] Angelsky, O.V.; Demianovsky, G. V.; Ushenko, A. G.; Burkovets, D. N. & Ushenko, Yu. A. (2004). Wavelet analysis of two-dimensional birefringence images of architectonics in biotissues for diagnosing pathological changes, *J. Biomed. Opt.*, V. 9, (2004), pp. 679-690
- [112] Angelsky, O.V.; Tomka, Yu.Ya.; Ushenko, A.G.; Ushenko, Ye.G. & Ushenko, Yu.A. (2005). Investigation of 2D Mueller matrix structure of biological tissues for pre-clinical diagnostics of their pathological states, *Journal of Physics D: Applied Physics*, V. 38(23), (2005), pp. 4227-4235
- [113] Cowin, S.C. (2000). How is a tissue built?, *J. Biomed. Eng.*, V. 122, (2000), pp. 553-568
- [114] Ushenko, A.G.; Pishak, V.P.; Yermolenko, S.B.; Pishak, O.V. & Burkovets, D.N. (1997). Laser measurements of crystal optical properties of blood-formed elements, *Proc. of SPIE* 1997, V. 3317, pp. 425-433

- [115] Ushenko, A.G. (1999). Laser diagnostics of biofractals, *Quantum Electronics*, V. 29(12), (1999), pp. 1078-1084
- [116] Angelsky, O.; Burkovets, D.; Pishak, V.; Ushenko, Yu. & Pishak, O. (2000). Polarization-correlation investigations of biotissue multifractal structures and their pathological changes diagnostics, *Laser Physics.*, V. 10, №5, (2000), pp. 1136-1142
- [117] Gori, F. (1998). Matrix treatment for partially polarized, partially coherent beams, *Opt. Lett.*, V. 23, (1998), pp. 241-243
- [118] Ushenko, Ye. G. (2005). Complex Degree of Mutual Polarization of Biotissue's Speckle-Images, *Ukr. J. Phys Opt.*, V. 6, No. 3, (2005), pp. 104-113
- [119] Ushenko, A.G. (1995). Polarization structure of laser scattering fields, *Optical Engineering*, V. 34(4), (1995), pp. 1088-1093
- [120] Ushenko, A.G. (2000). Polarization Structure of Biospeckles and the Depolarization of Laser Radiation, *Optics and Spectroscopy*, V. 89(4), (2000), pp. 597-601
- [121] Ushenko, A.G.; Ermolenko, S.B.; Burkovets, D.N. & Ushenko, Yu.A. (1999). Polarization Microstructure of Laser Radiation Scattered by Optically Active Biotissues, *Optics and Spectroscopy*, V. 87(3), (1999), pp. 434-439

# Digital Demodulation of Interferometric Signals

Tristan J. Tayag and R. Collins Watson  
*Texas Christian University*  
USA

## 1. Introduction

The marriage of optical sensing techniques with sophisticated digital signal processing has resulted in a myriad of practical metrology systems. Optical metrology systems offer many attractive measurement features. These systems are inherently non-contacting, non-destructive, and immune from electromagnetic interference. In addition, since light is used as the sensing probe, the measurement system is capable of high sensitivity, fine resolution (both spatial and temporal), and absolute calibration.

The leap from the vibration stabilized table of the laboratory to the harsh milieu of the factory floor has been a major challenge for optical metrology systems. However, on-going advances in digital processor speed and algorithmic complexity have in large part made this leap possible. Full-field optical inspection methods include holographic interferometry, speckle metrology, and interferometric computed tomography (Tayag & Bachim, 2010). These techniques have been applied to a wide range of applications in biological tissue characterization, the automotive industry, dentistry, the semiconductor industry (Pitt et al., 2003; Tayag et al., 2003; Tayag et al., 2010; Weber et al., 2004), and fiber optic and bulk optic characterization (Osten et al., 2010). Here, we focus our attention on digital demodulation applied to “point measurements” as opposed to full-field measurements.

In this chapter of *Modern Metrology Concerns*, we present recent digital signal processing techniques used in interferometry. Application areas range from biotechnology to industrial to military. Specific examples will be cited throughout the literature as well as from our own research. Our research in digital demodulation techniques include applications in myofibroblast cell contraction (Kern et al., 2003), ballistic shock characterization (Kumar et al., 2009), and down-hole oil well drilling. In addition, we present a new (unpublished) digital demodulation algorithm. The novelty of our technique lies in the frequency domain manipulation of the well-known phase-generated carrier modulation approach.

Section 2 contains background information on interferometry and practical modulation techniques. The Michelson, Mach-Zehnder, Sagnac, and Fizeau optical configurations will be described. Once the basic optical geometries are established, we describe the modulation and demodulation techniques which are needed in practice. These techniques convert the measured phase change into the desired displacement, velocity, or index change.

Our novel digital demodulation algorithm will be presented in Section 3. The application of this research is in sensing down-hole drill bit parameters for oil well drilling systems. We present a new demodulation algorithm for phase-generated carrier based interferometers. The theoretical background is presented along with simulation results.



Section 4 will follow with a summary and directions for future research.

## 2. Interferometry

Measurements made with interferometers are based on the interference pattern formed by two or more electromagnetic fields. In this discussion, we confine ourselves to the interference of two electromagnetic fields in the visible region of the spectrum. A stable fringe pattern results when the two interfering fields satisfy the following conditions:

- The two fields have the same optical wavelength.
- The two fields maintain a fixed phase relationship.
- Each of the two fields has its electric field vector linearly polarized.
- The two fields have collinear polarization vectors.

To achieve these conditions, the source may be a linearly polarized single mode (temporal and spatial) laser. The light from this laser is split into two beams known as the reference and signal beams. The output optical power resulting from the interference of these two fields takes the form of a raised cosine which varies as a function of the optical path length difference between the reference and signal beams (see Fig. 1).

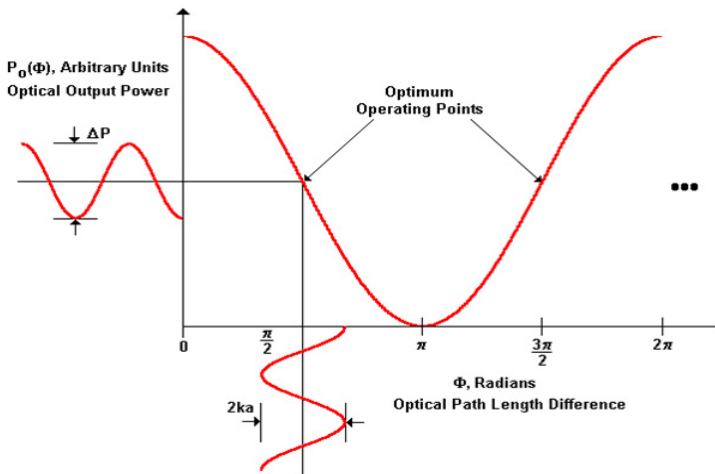


Fig. 1. Optical output power versus path length difference for an interferometer.

Practical metrology systems require fine resolving capabilities as well as large dynamic range. Fractional fringe interferometers are capable of resolving displacements that are many orders of magnitude less than the optical probe's wavelength. This high resolution is achieved when the interferometer is operated at quadrature, the point of maximum sensitivity. As shown in Fig. 1, operating the interferometer at quadrature yields large changes in the optical output power for small changes in the path length difference between the reference and signal beams. To achieve large dynamic range, interferometers may employ fringe counting techniques for changes in the optical path difference, which exceed half of the field's wavelength.

A number of two beam interferometric configurations have been developed over the past two centuries. In the next section, we will review four of these optical configurations.

## 2.1 Optical configurations

In 1881, the American physicist, Albert A. Michelson, developed the basic interferometric configuration shown in Fig. 2. Light from the optical source is split into the signal and reference beams with a beamsplitter. The reference beam reflects off a mirror and retraces its path. The signal beam reflects off the specular surface of the object whose displacement is to be measured. The two beams are recombined using the original beamsplitter. A portion of each of the reference and signal fields is incident on the detector. If the optical path length difference between the reference and signal fields is zero (or an integral multiple of wavelengths), the interference is constructive. If the optical path length difference is  $\pi$  (or an odd multiple of half wavelengths), the interference is destructive.

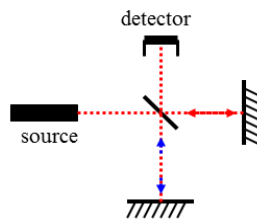


Fig. 2. Michelson optical configuration.

The Mach-Zehnder optical configuration shown in Fig. 3 provides flexibility over the Michelson configuration by using a second beamsplitter to combine the reference and signal beams. With the environment of the reference beam remaining constant, changes in the refractive index through which the signal arm passes may be measured. Because of its suitability for a “push-pull” arrangement, this configuration is used in optical modulator technology (Kaplan & Ruschin, 2000) as well as optical sensor technology (Porte et al., 1999).

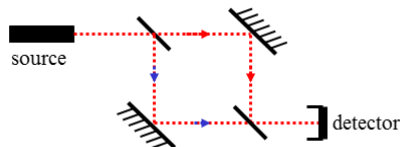


Fig. 3. Mach-Zehnder optical configuration.

The Sagnac or cyclic optical configuration is shown in Fig 4. This configuration is unique in that the two beams follow the same path around a closed circuit, but in opposite directions. This sensor measures the non-reciprocal phase changes that arise between light propagating in clockwise and counter-clockwise directions. In gyroscope sensing applications, the phase shift gives a measure of the rotation of the loop about its axis (Saida & Hotate, 1999; Barbour & Schmidt, 2001; Tselikov et al., 1998).

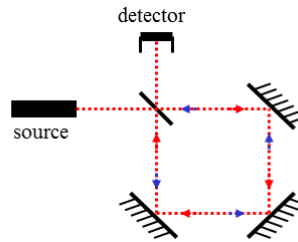


Fig. 4. Sagnac optical configuration.

The optical configuration shown in Fig. 5 is attributed to the French physicist, Hippolyte L. Fizeau. In this configuration, interference occurs between light reflected from the target and the partial reflection from the facet of the fiber probe. High stability (i.e., low phase drift) is a key advantage of this configuration since the fiber/target separation distance is typically very small to accommodate sufficient light coupling back into the fiber. Another advantage of this system is the common path travelled by the majority of the reference and signal beams. In the fiber optic embodiment shown in Fig. 5, this means that the single mode optical fiber does not have to be polarization preserving, since both beams will undergo the same polarization evolution as they travel through the fiber. Each of the interferometric configurations described in this section, encodes a displacement, index change, or rotation rate into a phase shift between two interfering light beams. Practical techniques for extracting the desired parameter from the phase shift information form the topics of the next section.

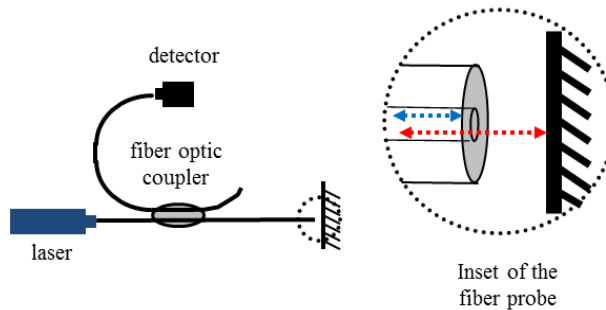


Fig. 5. Fizeau optical configuration.

## 2.2 Modulation and demodulation techniques

Interferometer modulation and demodulation techniques have been broadly classified into homodyne and heterodyne approaches. Homodyne modulation techniques refer to those systems where the two interfering beams of light are of the same optical frequency. The homodyne techniques may be further sub-divided into active homodyne and passive homodyne. In the early days of interferometry, the scientist would wait until the interferometer drifted into the quadrature operating point before starting data collection. Today, active feedback control is used to adjust the reference path and maintain the interferometer at quadrature. Digital signal processing of the modulated signal is applied to compute the feedback control signal.

In passive homodyne modulation, techniques are used to generate output signals which after digital processing eliminate the signal fading that occurs as the interferometer drifts from the quadrature condition. In this way, no feedback control loop is needed to maintain the optimum operating point.

Heterodyne modulation techniques refer to those systems where the two interfering beams have differing optical frequencies. A frequency shift in one of the beams may be accomplished by acousto-optic frequency modulation (Culshaw & Giles, 1982), frequency ramping the optical source (Jackson et al., 1982), or piezoelectric-induced path length changes (Jackson et al., 1980; Cole et al., 1982). These heterodyne modulation approaches have been categorized as true-heterodyne, pseudo-heterodyne, and synthetic-heterodyne techniques.

In practical application, the boundaries between the homodyne and heterodyne approaches as well as their sub-divisions become blurred. In the remainder of this section, we describe 3 digital demodulation methods for interferometric sensors.

The phase-generated carrier demodulation technique was first proposed by Dandridge, Tveten, and Giallorenzi in 1982 (Dandridge et al., 1982). It is a homodyne demodulation technique which eliminates signal fading as slow environmental perturbations drift the interferometer's path length difference away from quadrature. In this technique, a sinusoidal modulation with known frequency and amplitude is imposed on the interferometer's phase difference. Detection and mathematical manipulation of interferometer signal allow the desired phase shift to be separated from environmental perturbations. The signal processing involves detection of the signal amplitudes at the carrier's fundamental and second harmonic frequencies. The time derivative of each of these signals is computed and they are cross-multiplied with the original signals. Integration of the difference of these signals results in the desired phase change. We have analyzed the quantum-noise-limited sensitivity of interferometers using a phase-generated carrier modulation scheme (Tayag, 2002).

Dandridge et al. first demonstrated this phase-generated carrier demodulation technique using analog electronics. Since that time, a commercially-available demodulation system based on the phase-generated carrier has come on the market (Bush et al., 1996; Cekorich et al., 1997; Davis et al., 1998). This system is based on real-time digital processing of the interferometer's signal in the time domain. We have used this demodulation system to characterize the contraction of myofibroblast cells within a collagen matrix (Kern et al., 2003). In Section 3 of this chapter, we present a new demodulation algorithm whose kernel is based on a frequency domain analysis of the interferometer's phase-generated carrier signal.

The J1...J4 demodulation method (Sudarshanam & Srivivasan, 1989) is a passive homodyne technique which requires no feedback for stabilization of the interferometer at quadrature. This technique is applicable to measuring the amplitude of sinusoidally vibrating structures. In this approach, the photodetected interferometer output is expanded using a Fourier-Bessel series. In fact, the "J1...J4" name of this method refers to the  $J_n(x)$  Bessel functions of the first kind of order  $n$  in the Fourier-Bessel expansion. Jin et al. (Jin et al., 1991; Jin et al., 1992) modified this method to extend the desired phase modulation index. Marcal et al. (Marcal et al., 2007) have used this digital demodulation technique with a Michelson interferometric configuration to characterize the frequency response of a novel piezoelectric actuator. We have used a similar digital demodulation algorithm to characterize

microelectromechanical system (MEMS) structures (Pitt et al., 2003). Our fiber optical configuration for this system was based on the Fizeau interferometer.

A final digital demodulation technique we will review is based on the use of a 3x3 fiber optic coupler in the optical system (Kwon et al., 1999; Sheem, 1981; Sheem et al., 1982). Again, the demodulation method is classified as a passive homodyne technique. The optical configuration is shown in Fig. 6 and is basically a Mach-Zehnder optical configuration, but uses a 3x3 fiber coupler to combine the beams. Koo, Tveten, and Dandridge (Koo et al., 1982) were the first to introduce this passive stabilization technique for fiber optic interferometers. Todd et al. (Todd et al., 1999) later applied digital demodulation techniques to Koo's 3x3 fiber coupler-based system.

The unbalanced path lengths and 3x3 coupler of the Mach-Zehnder interferometer shown in Fig. 6 are necessary for the interferometer to function as a wavelength discriminating sensor. The outputs from the 3 photodetectors are shifted from each other by  $2\pi/3$  radians. With *a priori* knowledge of the coupler and detector characteristics, the post-detection digital processing can extract the interferometer's phase shift corresponding to a wavelength shift in the optical signal. This interferometric wavelength discriminator has been used to interrogate and multiplex a network of fiber Bragg grating sensors (Johnson et al., 2000) and as a velocity sensor which measures the Doppler shift in the light reflected from a moving surface (Fabiny & Kersey, 1997). We have also used this system as a Doppler velocimeter to measure the motion of armored plates experiencing ballistic shock (Kumar et al., 2009).

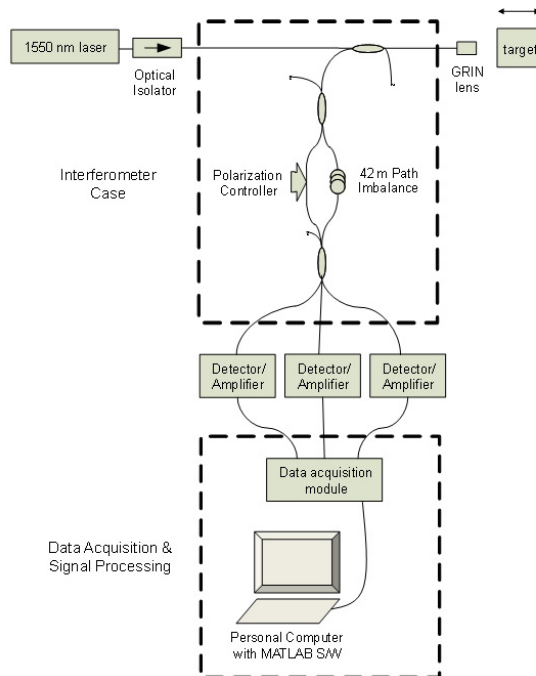


Fig. 6. Interferometric velocimeter based on a 3x3 fiber coupler and homodyne demodulation technique (Kumar et al., 2009).

### 2.3 Summary

In this section, we have reviewed the classical optical configurations and modulation/demodulation techniques of interferometric systems. There exists a close relationship between the optical configuration and the modulation technique. Our focus has been on those systems which exploit digital processing in their demodulation methods. It is interesting to note that all the interferometric systems we have presented are based on passive homodyne demodulation. This is not surprising, since the digital processor and its associated demodulation algorithm are replacing the optical hardware (acousto-optic modulators or swept frequency lasers and their drive electronics) and electronic hardware (piezoelectric transducers, analog differentiators, integrators, and lock-in amplifiers) necessary in active homodyne and heterodyne systems. In this way, the overall interferometric system is greatly simplified with the burden of complexity falling on the software. This results in a more robust and less expensive metrology system.

### 3. Frequency domain demodulation of a phase-generated carrier

Since the time and frequency domains represent different ways of describing the same signal, it is reasonable to consider an algorithm based on frequency domain analysis as an alternative to the phase-generated carrier time domain algorithm discussed in the previous section. In this section, we describe the kernel of a novel algorithm for the demodulation of an interferometric signal based on a phase-generated carrier modulation scheme. Following this section, practical implementation of the algorithm is discussed with simulation results verifying the algorithm. Finally, the frequency domain demodulation algorithm is summarized with suggestions for further research.

#### 3.1 Algorithmic development

Two-beam interferometers produce an optical interference pattern, which varies with optical path length difference,  $R(t)$ , between a signal arm and a reference arm. It is well-known that the optical power,  $P(t)$ , of this interference pattern, takes the form

$$P(t) = P_{dc} + P_{ac} \cos[R(t)],$$

where  $P_{dc}$  is the dc optical power and  $P_{ac}$  is the 0-to-peak modulation index of the interference pattern. This optical power is incident on a photodetector for conversion to an electronic signal. Since PIN photodiodes are characterized by a responsivity, which is constant with respect to the incident optical power, the photo-generated current takes the same mathematical form of a raised cosine. If the electronic signal from the receiver is a voltage, then this signal takes the form

$$V(t) = V_{dc} + V_{ac} \cos[R(t)],$$

where  $V_{dc}$  is the dc voltage level and  $V_{ac}$  is the 0-to-peak modulation index of the interference pattern. The objective of the demodulator is to sample this electronic signal, process the signal samples in real-time to determine the phase  $R(t)$ , and produce an analog signal output, which is proportional to the phase  $R(t)$ .

The demodulation algorithm is based on a phase generated carrier modulation technique. A key assumption is that a sinusoidal carrier signal is imposed on the interferometer's optical signal such that the interference pattern becomes

$$V(t) = V_{dc} + V_{ac} \cos[R(t) + M \sin(2\pi Ft + W)], \quad (1)$$

where  $M$  is the phase index of the sinusoidal carrier,  $F$  is the frequency of the carrier, and  $W$  is the phase of the carrier. Since the sinusoidal carrier appears as a phase term in the interferometer's cosinusoidal interference pattern, this form of modulation is referred to as a "phase-generated carrier."

The output voltage of Eqn. (1) describes the interferometer's output signal. Let us rewrite this equation with the angular frequency of the carrier signal defined by  $\omega_o \equiv 2\pi F$  rad/s:

$$V(t) = V_{dc} + V_{ac} \cos[R + M \sin(\omega_o t + W)]. \quad (2)$$

We may set our sinusoidal carrier signal such that  $M = \pi$  rad and  $W = 0$  rad. Then Eqn. (2) becomes

$$V(t) = V_{dc} + V_{ac} \cos[R + \pi \sin(\omega_o t)]. \quad (3)$$

Trigonometric expansion of the cosine term yields

$$V(t) = V_{dc} + V_{ac} \{ \cos[R] \cdot \cos[\pi \sin(\omega_o t)] - \sin[R] \cdot \sin[\pi \sin(\omega_o t)] \}. \quad (4)$$

Now, note the Bessel expansions

$$\cos[\pi \sin(\omega_o t)] = J_0(\pi) + 2 \sum_{n=1}^{\infty} J_{2n}(\pi) \cos[(2n) \cdot \omega_o t] \quad \text{and} \quad (5)$$

$$\sin[\pi \sin(\omega_o t)] = 2 \sum_{m=1}^{\infty} J_{2m-1}(\pi) \sin[(2m-1) \cdot \omega_o t], \quad (6)$$

where  $J_n(x)$  is a Bessel function of the first kind of order  $n$ . Substitution of Eqns. (5) and (6) into Eqn. (4) reveals the harmonic content of the interferometric signal in the frequency domain as

$$\begin{aligned} V(t) = & [V_{dc} + V_{ac} \cos(R) J_0(\pi)] \\ & - 2V_{ac} \sin(R) J_1(\pi) \sin(\omega_o t) \\ & + 2V_{ac} \cos(R) J_2(\pi) \cos(2\omega_o t) \\ & - 2V_{ac} \sin(R) \sum_{m=2}^{\infty} J_{2m-1}(\pi) \sin[(2m-1) \cdot \omega_o t] \\ & + 2V_{ac} \cos(R) \sum_{n=2}^{\infty} J_{2n}(\pi) \cos[(2n) \cdot \omega_o t]. \end{aligned} \quad (7)$$

The first and second harmonics of  $V(t)$  are

$$V_{\omega_0}(t) = -2V_{ac} \sin(R) J_1(\pi) \sin(\omega_0 t) \quad \text{and} \quad (8)$$

$$V_{2\omega_0}(t) = 2V_{ac} \cos(R) J_2(\pi) \cos(2\omega_0 t). \quad (9)$$

These two frequency components contain sufficient information to determine the desired phase,  $R$ . By taking the ratio of the amplitudes of these frequencies, the dependence on the interference pattern's modulation index  $V_{ac}$  is removed. The Bessel functions are known constants, which may be removed from the ratio as follows:

$$\frac{J_2(\pi) [V_{\omega_0}(t) \text{ amplitude}]}{J_1(\pi) [V_{2\omega_0}(t) \text{ amplitude}]} = \frac{\sin(R)}{\cos(R)} = \tan(R). \quad (10)$$

The phase  $R$  may be computed from this quotient over an entire  $2\pi$  rad phase range using an inverse tangent approximation. This forms the kernel of the frequency domain demodulation algorithm. A fringe counting approach similar to that described by Cekorich (Cekorich, 1999), may be used to extend the range of demodulation. In the next section, we describe practical DSP techniques to implement the frequency domain demodulation algorithm.

### 3.2 Simulation results

We have verified the frequency domain demodulation algorithm through computer simulation. In this section, we present those simulation results and describe our DSP implementation approach. The frequency spectrum of the interferometer's output signal is shown in Fig. 7. The magnitude spectra are shown for the five different values of the desired phase  $R$ . The periodic nature of the time domain signals correctly results in a set of discrete, harmonically-related sinusoids in the frequency domain. The open circles in Fig. 7 indicate the fundamental and second harmonic frequencies which are oscillating at  $\omega_0$  and  $2\omega_0$ , respectively. Note the consistency between the magnitude spectra of Fig. 8 and Eqns. (8) and (9). Specifically, the zeros are apparent at  $\omega_0$  for  $R=0$  and  $\pi$  rad and at  $2\omega_0$  for  $R=\pi/2$  rad. As described by Eqn. (10), these two frequency components will be used to demodulate the interferometer.

A significant aspect of the frequency domain algorithm is the conversion of the sampled time domain signal into a frequency domain representation. Computation of the frequency spectrum is conveniently performed via the well-known fast Fourier transform (FFT) algorithm. Since only 2 spectral components are required for this algorithm, the Goertzel algorithm provides a slightly more computationally efficient approach. We implemented both algorithms, but only present the results for the FFT algorithm here since the Goertzel results are equivalent.

The FFT algorithm results in complex-valued signals at each frequency bin. The magnitudes of these complex signals are shown in Fig. 7. Fig. 8 illustrates how the magnitudes of the



fundamental and second harmonics vary with the phase,  $R$ . These plots are consistent with the mathematical form given in Eqns. (8) and (9). The difference in peak amplitudes results from the Bessel function weighting factors, i.e.  $J_1(\pi) = -0.5692$  and  $J_2(\pi) = 0.9709$ . Note that the magnitude spectrum contains only positive values. Therefore, the magnitude spectrum alone is insufficient to unambiguously determine  $R$ . To resolve  $R$  over the full  $2\pi$  phase interval, the phase spectrum is also needed. Fig. 9 shows the phases of the fundamental and second harmonics as a function of  $R$ .

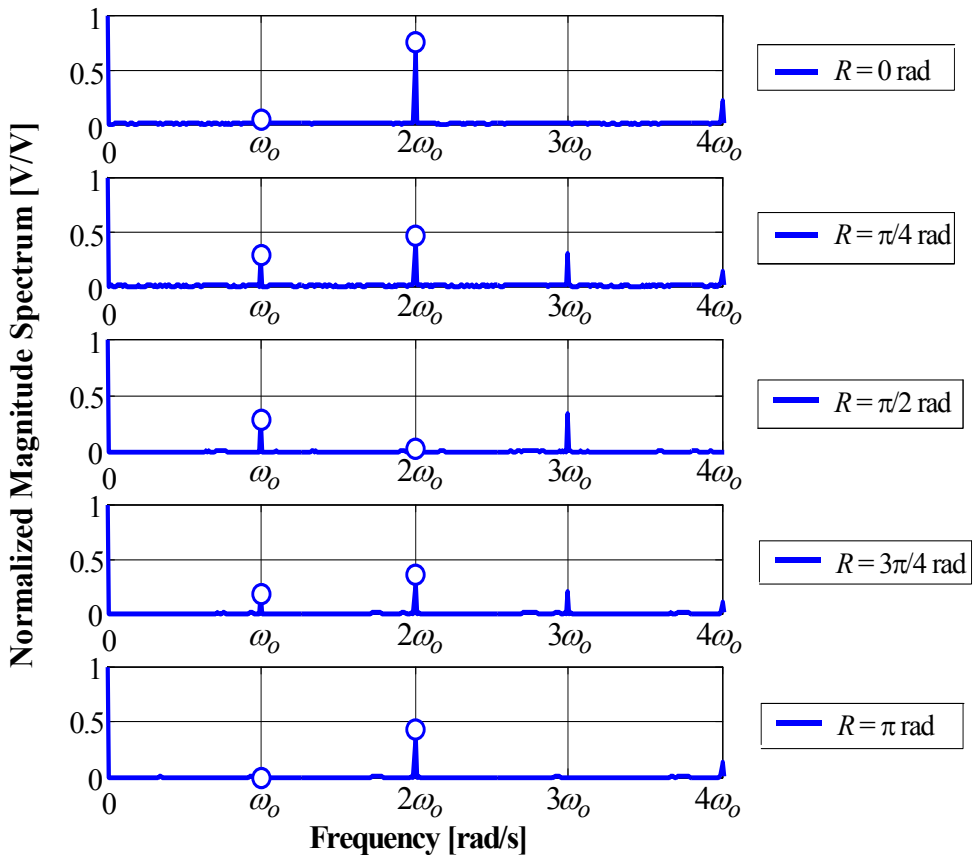


Fig. 7. Normalized frequency spectra of the interferometer signal described by Eqn. (7) at five different values of the phase  $R$ . The open circles represent the values of the fundamental and second harmonic frequency components.

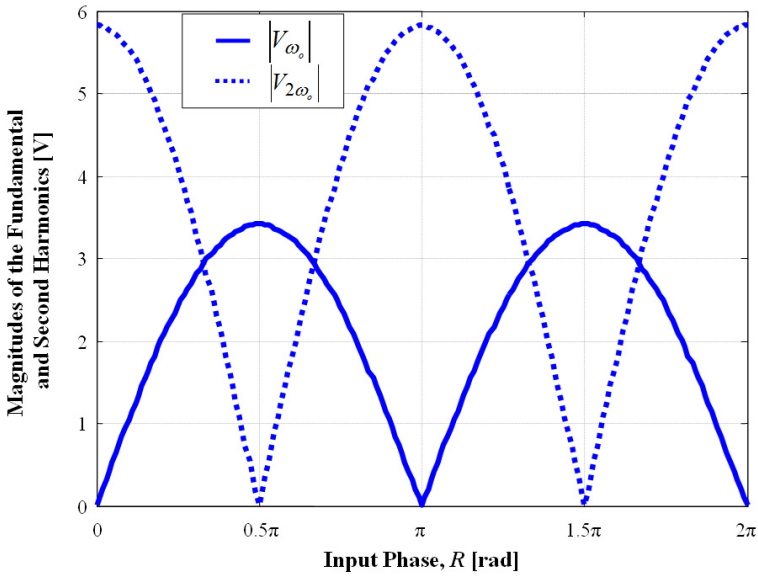


Fig. 8. Magnitudes of the fundamental and second harmonic frequencies as a function of the phase  $R$ .

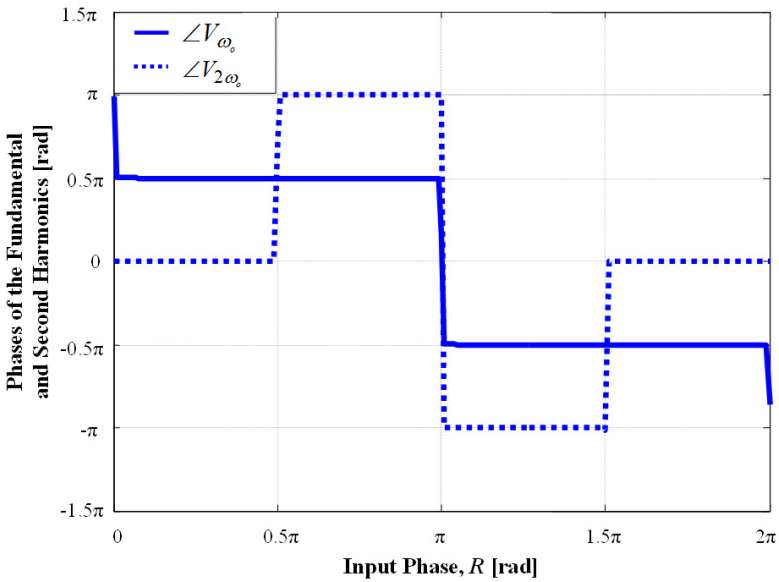


Fig. 9. Phases of the fundamental and second harmonic frequencies as a function of the phase  $R$ .

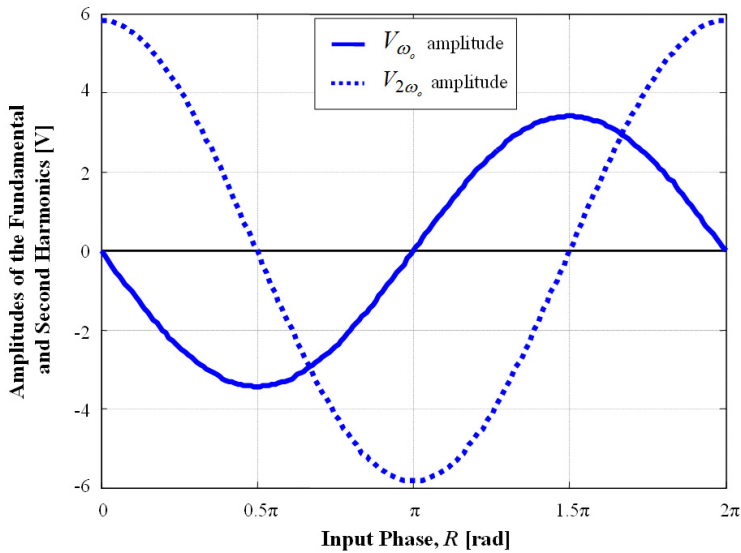


Fig. 10. Reconstructed amplitudes of the fundamental and second harmonic frequencies as a function of the phase  $R$ .

A key to reconstructing the amplitudes of the fundamental and second harmonic frequencies is to utilize both the magnitude and phase information provided by the signal's frequency spectrum. This information shown in Figs. 8 and 9 may be used to determine the amplitudes of Eqns. (8) and (9). Fig. 10 shows the amplitudes of the fundamental and second harmonics reconstructed from the data of Figs. 8 and 9. Notice that the amplitude of the fundamental harmonic varies sinusoidally as a function of  $R$  and the second harmonic varies cosinusoidally as a function of  $R$ . This data applied to Eqn. (10) may be used with an inverse tangent approximation to determine the desired phase  $R$  over a full  $2\pi$  phase interval. This forms the kernel of the frequency domain demodulation algorithm. We have incorporated a fringe counting routine and verified the frequency domain demodulation algorithm through computer simulation.

In this section, we have verified the theoretical development of the previous section through computer simulation and described how to implement the algorithm in practice. A key step in the kernel of the frequency domain algorithm is use of both the magnitude and phase components of the frequency spectrum.

### 3.3 Discussion

We have presented a frequency domain approach to the demodulation of an interferometer, which uses a phased-generated carrier modulation scheme. Through simulation analysis, we have shown that the frequency domain algorithm produces correct demodulation results by assuming fixed operating parameters (i.e.,  $M=\pi$  and  $W=0$ ). These are the same operating conditions as used in commercially-available time-domain demodulators.

Unfortunately, the frequency domain algorithm suffers from the same limitations as the time domain algorithm with regard to the carrier's modulation depth and phase. Drift of the carrier's modulation depth from  $M = \pi$  produces a change in the amplitudes of the fundamental and second harmonic frequencies. This dependence manifests itself in the Bessel functions in Eqn. (7). If the value of this drift is unknown, it cannot be corrected in the computation of Eqn. (10) and errant demodulation results. Similarly, a drift of the carrier's modulation phase from  $W = 0$  produces an irreparable change in the amplitudes of the harmonic frequencies.

As a consequence of these limitations, the frequency domain algorithm (like the time domain algorithm) requires careful initialization of the carrier's modulation depth and phase. Practical implementation of interferometric systems often use hardware, such as piezoelectric transducers, whose modulation depth drifts over time. Therefore, closed-loop active feedback is necessary to maintain the requisite modulation depth at  $M = \pi$  rad.

### 3.4 Summary

We have presented the kernel of an algorithm for the digital demodulation of an interferometer based on the phase-generated carrier modulation scheme. This algorithm exploits both the magnitude and phase information in its frequency domain manipulation of the signal. The algorithm suffers from requirement that active feedback control is needed.

## 4. Conclusion

In this chapter, we have reviewed digital demodulation algorithms for interferometric metrology systems. Digital demodulation has the potential for producing a more robust and less costly sensor system. In general, the savings in optical and electronic hardware is traded-off with complex software algorithms. Consider the information shown in Table 1 of the specific digital demodulation techniques presented in this chapter.

Note that each of the digital demodulation methods presented in Table 1 is based on a homodyne approach. The heterodyne techniques require additional optical and electronic hardware to shift the frequency of one of the interfering optical beams. A key advantage of digital demodulation is a significant reduction in cost and hardware system complexity.

We have also presented a new demodulation algorithm for the well-known phase-generated carrier modulation scheme. Our algorithm is based on a frequency domain approach and requires both the magnitude and phase information present in the signal. This demodulation scheme suffers from the need for active feedback control. One possible approach to ameliorate this limitation is to investigate the use of a more sophisticated modulation signal. In the demodulation algorithm presented in Section 3, the carrier signal is a pure sinusoid and the desired output is the fixed phase  $R$ . We have studied interferometers, which use the phase-generated carrier modulation scheme and detect the amplitude of a *sinusoidally* varying phase  $R(t)$ . These systems produce harmonics, which we detect and then compute their ratio to remove the dependence on unknown quantities.

Demodulation Technique	Phase-generated carrier (time domain)	Phase-generated carrier (frequency domain)	J1 ... J4	3x3 fiber coupler
Application	Biological cell contraction measurement	Down-hole oil well drilling bit parameter measurement	MEMS vibrating amplitude measurement	Ballistic shock (velocity) measurement
Homodyne/Heterodyne	Homodyne	Homodyne	Homodyne	Homodyne
Passive/Active	Active	Active	Passive	Passive
Real-time/Post-processing	Real-time processing	Real-time processing	Real-time processing	Post-processing

Table 1. Comparison of digital demodulation techniques presented in this chapter.

It may be possible to design a carrier signal, which produces an appropriate set of harmonics that can be mathematically manipulated to cancel out any variations in  $M$  and  $W$ , yet still be solvable for the fixed phase  $R$ . The trade-off would be a true open-loop system, which would not require stringent initialization procedures, at the expense of a sophisticated carrier modulation signal. Further investigation into the viability of this approach is suggested.

## 5. Acknowledgment

The authors would like to recognize funding support from Weatherford International, Inc. for research related to the phase-generated carrier demodulation scheme presented in Sec. 3.

## 6. References

- Barbour, N. & Schmidt, G., Inertial sensor technology trends. *IEEE Sensors Journal*, vol. 1, no. 4, (December 2001), pp. 332-339
- Bush, J., Davis, C. A., McNair, F., Cekorich, A., & Bostick, J., Low cost fiber optic interferometric sensors. *Proceed. of the SPIE - Second Pacific Northwest Fiber Optic Sensor Workshop*, vol. 2872, (May 1996), pp. 1-12.
- Cekorich, A., Bush, J., & Kirkendall, C., Multi-channel interferometric demodulator. *Proceed. of the SPIE - Third Pacific Northwest Fiber Optic Sensor Workshop*, vol. 3180, (May 1997), pp. 1-11.
- Cekorich, A., Demodulator for interferometric sensors. *Proceed. of the SPIE*, vol. 3860, (December 1999), pp. 338-347.
- Cole, J. H., Danver, B. A., & Bucaro, J. A., Synthetic-heterodyne interferometric demodulation. *IEEE J. Quantum Electron.*, vol. QE-18, no. 4, (April 1982), pp. 694-697.
- Culshaw, B. & Giles, I. P., Frequency modulated heterodyne optical fiber Sagnac interferometer. *IEEE J. Quantum Electron.*, vol. QE-18, no. 4, (April 1982), pp. 690-693).

- Dandridge, A., Tveten, A. B., & Giallorenzi, T. G., Homodyne demodulation scheme for fiber optic sensors using phase generated carrier. *IEEE J. Quantum Electron.*, vol. QE-18, no. 10, (October 1982), pp. 1647-1653.
- Davis, P. G., Bush, I. J., & Maurer, G. S., Fiber optic displacement sensor. *Proceed. of the SPIE – Fourth Pacific Northwest Fiber Optic Sensor Workshop*, vol. 3489, (September 1998), pp. 18-22.
- Fabiny, L. & Kersey, A. D., Interferometric fiber-optic Doppler velocimeter with high-dynamic range. *IEEE Photonics Technol. Lett.*, vol. 9, no. 1, (January 1997), pp. 79-81.
- Jackson, D. A., Priest, R., Dandridge, A., & Tveten, A. B., Elimination of drift in a single-mode optical fiber interferometer using a piezoelectrically stretched coiled fiber. *Appl. Opt.*, vol. 19, no. 17, (September 1980), pp. 2926-2929.
- Jackson, D. A., Kersey, A. D., Corke, M., & Jones, J. D. C., Pseudoheterodyne detection scheme for optical interferometers. *Electron. Lett.*, vol. 18, no. 25, (December 1982), pp. 1081-1083.
- Jin, W., Zhang, L. M., Uttamchandani, D., & Culshaw, B., Modified J1...J4 method for linear readout of dynamic phase changes in a fiber-optic homodyne interferometer. *Appl. Opt.*, vol. 30, no. 31, (November 1991), pp. 4496-4499.
- Jin, W., Walsh, D., Uttamchandani, D., & Culshaw, B., (1992). A digital technique for passive demodulation in a fiber optic homodyne interferometer. *1<sup>st</sup> European Conf. On Smart Structures and Materials*, Glasgow, UK, pp. 57-60.
- Johnson, G. A., Todd, M. D., Althouse, B. L., & Chang, C. C., Fiber Bragg grating interrogation and multiplexing with a 3x3 coupler and a scanning filter. *J. Lightwave Technol.*, vol. 18, no. 8, (August 2000), pp. 1101-1105.
- Kaplan, A. & Ruschin, S., Layout for polarization insensitive modulation in LiNbO<sub>3</sub> waveguides. *IEEE J. Selected Topics in Quantum Electron.*, vol. 6, no. 1, (January 2000), pp. 83-97.
- Kern, J., Dimitrijevic, S. D., Tayag, T. J., Pitt, B. D., Summers, T. H., & Davis, C. G., An optical system to characterize the gross contractile response of a tissue-equivalent collagen matrix. *Proceed. of the SPIE - Optical Technologies to Solve Problems in Tissue Engineering*, vol. 4961, (January 2003), pp. 244-249.
- Koo, K. P., Tveten, A. B., & Dandridge, A., Passive stabilization scheme for fiber interferometers using (3x3) fiber directional couplers. *Appl. Phys. Lett.*, vol. 41, no. 7, (October 1982), pp. 616-618.
- Kumar, P., Thomas, A., Weis, R. S., & Tayag, T. J., Digital processing of an interferometric velocimeter for ballistic shock measurement. *Proceed. of the SPIE – Optical Inspection and Metrology for Non-Optics Industries*, vol. 7432, (2009).
- Kwon, I. B., Kim, C. G., & Hong, C. S., A digital signal processing algorithm for structural strain measurement by a 3x3 passive demodulated fiber optic interferometer sensor. *Smart Materials & Structures*, vol. 8, (1999), pp. 433-440.
- Marcal, L. A. P., Leao, J. V. F., Nader, G., Higuti, R. T., Kitano, C., & Silva, E. C. N., Analysis of linearity and frequency response of a novel piezoelectric flexensional actuator using a homodyne interferometer and the J1-J4 method. *IEEE Trans. on Instrumentation and Measurement*, vol. 56, no. 3, (June 2007), pp. 954-961.
- Osten, W., Garbusi, E., Fleischle, D., Lyda, W., Pruss, C., Reichle, R., & Falldorf, C., Optical metrology – from the laboratory to the real world. *Proceed. of the SPIE – Speckle 2010: Optical Metrology*, vol. 7387, (2010).

- Pitt, B. D., Tayag, T. J., & Nelson, M. L., Digital demodulation of an interferometer for the characterization of vibrating microstructures. *Proceed. of the SPIE - Advance Semiconductor Characterization Tech. for Optics, Semiconductors, and Nanotechnologies*, vol. 5188, (August 2003), pp. 61-70.
- Porte, H., Gorel, V., Kiryenko, S., Goedgebuer, J.-P., Daniau, W., & Blind, P., Imbalanced Mach-Zehnder interferometer integrated in micromachined silicon substrate for pressure sensor. *J. Lightwave Technol.*, vol. 17, no. 2, (February 1999), pp. 229-233.
- Saida, T. & Hotate, K., General formula describing drift of inteferometer fiber-optic gyro due to Faraday effect: Reduction of the drift in twin-depo-I-FOG. *J. Lightwave Technol.*, vol. 17, no. 2, (February 1999), pp. 222-228.
- Sheem, S. K., Optical fiber interferometers with [3x3] directional couplers: Analysis. *J. Appl. Phys.*, vol. 52, no. 6, (June 1981), pp. 3865-3872.
- Sheem, S. K., Giallorenzi, T. G., & Koo, K., Optical techniques to solve the signal fading problem in fiber interferometers. *Appl. Opt.*, vol. 21, no. 4, (February 1982), pp. 689-693.
- Sudarshanam, V. S. & Srivivasan, K., Linear readout of dynamic phase change in a fiber-optic homodyne interferometer. *Opt. Lett.*, vol. 14, no. 2, (January 1989), pp. 140-142.
- Tayag, T. J., Quantum-noise-limited sensitivity of an interferometer using a phase generated carrier demodulation scheme. *Opt. Eng. Lett.*, vol. 41, no. 2, (February 2002), pp. 276-277.
- Tayag, T. J., Kolesar, E. S., Pitt, B. D., Hoon, K. S., Marchetti, J., & Jafri, I. H., An optical fiber interferometer for measuring the *in situ* deflection characteristics of MEMS structures. *Opt. Engineer.*, vol. 42, no. 1, (January 2003), pp. 105-111.
- Tayag, T. J. & Bachim, B., Simulation of an interferometric computed tomography system for intraocular lenses. *Proceed. of the SPIE - Interferometry XV: Applications*, vol. 7791, (2010).
- Tayag, T. J., Htun, T., & Kolesar, E. S., Integration of a low-cost fiber interferometer with a MEMS probe station. *Proceed. of the SPIE - Interferometry XV: Applications*, vol. 7791, (2010).
- Todd, M. D., Johnson, G. A., & Chang, C. C., Passive, light intensity-independent interferometric method for fibre Bragg grating interrogation. *Electron. Lett.*, vol. 35, no. 22, (October 1999), pp. 1970-1971.
- Tselikov, A., de Arruda, J. U., & Blake J., Zero-crossing demodulator for open-loop Sagnac interferometers. *J. Lightwave Technol.*, vol. 16, no. 9, (September 1998), pp. 1613-1619.
- Weber, R. A., Tayag, T. J., & Shannon, L. J., Digital demodulation algorithm for the interferometric characterization of RF MEMS structures. *Proceed. of the SPIE - Interferometry XII - Techniques and Analysis*, vol. 5531, (August 2004), pp. 315-322.

# Attosecond Metrology and Spectroscopy

Martin Schultze and Reinhard Kienberger

*Ludwig-Maximilians-Universität and Technische Universität München  
Germany*

## 1. Introduction

Time resolved investigations of electron dynamics in a variety of sample systems as atoms, molecules and solids establish the spearhead of modern ultrafast science. The technology that enables scientists to track and control these processes is referred to as “attosecond metrology” and is associated in symbiosis with ultrashort laser source development.

The origin of attosecond metrology is linked to the advent of ultrashort mode locked laser sources - femtosecond lasers, chirped pulse amplification, carrier-envelope-phase stabilization - and the technological aspects of this development that rendered experiments with a temporal resolution of tens of attoseconds possible. The high-order harmonic-generation process yields laser pulses in the wavelength region around 10 nm with durations below 100 attoseconds, at the same time being the shortest controllable signals. We introduce the different spectroscopic techniques that make use of this unprecedented temporal resolution. While the “attosecond transient recorder” constitutes an oscilloscope for electromagnetic fields oscillating at visible frequencies, “attosecond tunneling spectroscopy” as well as “attosecond transient absorption” provides real-time insight into the dynamic processes of electronic rearrangements in matter.

Looking at the evolution of the development of ultrashort pulses, it is remarkable that optical communication stood behind RF communication and electronics for more than a century. Light did not have to be “invented”, it was there – due to our ability to see electromagnetic radiation at wavelengths between 380 and 480 nm. Especially since man was able to control fire, light was used as a signal for information transport over wide distances: lighthouses, smoke signals, nautic flags; using optical telegraphy with movable signal arms, the 225 kilometers between Metz and Mainz could be covered in a few minutes. Just after the invention of the light bulb by Thomas Alva Edison and Heinrich Göbel, in 1886 Heinrich Hertz proved the radio waves which had been postulated by James Clerk Maxwell in 1864. Since switching rates for light bulbs were considerably smaller than the ones for radio waves, and since there were no transmission systems for light with low losses, within short time signal communication via electric cables and RF signals was established. The spectacular advances in the development of optical fibers with extremely low losses (Nobel prize 2009 for Charles Kao) as well as the development of semiconductor-based incoherent and coherent light sources (LEDs and diode lasers) lead to the comeback and triumph of optical communication at the end of the laser century. These sources made possible to generate sufficiently short pulses to achieve extremely high switching and transmission rates.



The needs in optical communications have motivated the development of light sources with short and ultrashort pulse duration significantly. Scientists from atomic, molecular and optical physics made significant contributions and profited from this development. After breaking the picosecond border in pulse duration, an own field, ultrafast physics evolved. Light pulses of ultrashort duration are a key tool to investigate processes in microcosm. Ahmed Zewail was awarded the Nobel prize 1999 for femtosecond chemistry (Anon 1994). The laws of quantum optics dictate the duration at which processes in atoms, molecules and solids take place (Ferenc Krausz 2009). Electronic transitions in atoms, for example, can happen on a timescale of femtoseconds or attoseconds, where an attosecond (10<sup>-18</sup> s) compares to one second approximately like a second to the age of the universe.

From photography we know that an image of a moving object can only be unblurred if the exposure time is significantly shorter than the duration of the movement. The time-resolution of the electronic processes mentioned above therefore needs pulse durations of less than one femtosecond. That's not so easy, since nature does not give us anything for free: Electromagnetic pulses, hence light pulses consist of many waves of different wavelengths which add up in a way that a temporarily confined signal – the pulse – is formed. The envelope within which the electric field oscillates with the carrier frequency defines the duration of the pulse. In order that a pulse can propagate in space, the carrier has to undergo at least one oscillation cycle, which sets a lower limit for the pulse duration. At 750 nm, which is a common wavelength for ultrafast lasers, one oscillation cycle lasts about 2.5 fs. Amplified laser pulses at this central wavelength can be made as short as 3.3 fs, today (A. L. Cavalieri, E. Goulielmakis, et al. 2007). There is not much to be improved in terms of pulse duration at this wavelength. To achieve even shorter pulse durations, shorter wavelengths in the extreme ultraviolet (XUV) or x-ray regime are needed. Unfortunately, nature doesn't help us here, again. At these wavelengths no materials exist that could be used as laser media, directly. Besides demanding and expensive methods like free-electron-lasers (FEL), which are big facilities with hundreds of meters extension, coherent light can be produced by a non-linear wavelength conversion from laser pulses in the visible range. The generation of high-order harmonics (McPherson et al. 1987) (Ferry et al. 1988) provides the possibility to produce attosecond pulses (Paul et al. 2001) (Christov et al. 1997). Before we concentrate on this method, let's have a look at the formation of ultrashort pulses comprising barely more than one oscillation cycle of the driver wave.

## 2. State-of-the-art

Laser pulses can be very short if their spectrum is extremely broad. At the high frequency of visible light in the range of Petahertz, it is sufficient to have a spectral width of a few Terahertz to achieve pulse durations of less than one picosecond. This can be realized at very different ways. Very common are quality switched (Q-switched) lasers, which are equipped with a device in the laser resonator that influences – actively or passively – the finesse of the resonator and therefore generates pulsed radiation. In order to achieve even shorter pulses, a number of technological approaches have been realized, making different wavelengths in the laser resonator oscillate synchronously. If the oscillations of the different longitudinal resonator modes, that are supported by the gain medium, have a constant phase relation (i.e. mode locking), a short pulse is created. The more resonator modes couple, the shorter is the pulse.

After dye lasers had dominated short pulse technology, today nearly all practically all short pulse lasers are based on solids, i.e. glasses and crystals, as the gain medium. In the visible and neighboring infrared range, sapphire doped with titanium ions has prevailed. Ti:Sa lasers have two main advantages – besides being relatively easy to handle: They can be pumped by commercially available, frequency doubled infrared lasers and have a comparably broad amplification bandwidth. To make use of this bandwidth as many modes in the resonator as possible have to be mode-locked and the propagation of a short pulse has to be favored with respect to continuous radiation. Aside from optical components that compensate for dispersion in the resonator, this is mainly made possible by saturable absorbers or concepts based on the optical Kerr effect.

Saturable absorbers are semiconductors with the special characteristic to escalate reflectivity when incoming radiation exceeds a certain threshold intensity. Such a component improves the finesse of a resonator selectively for short pulses and enables mode locking. Another common technology, which can be combined with saturable absorbers, is Kerr-lens mode-locking which makes use of the intensity dependence of the refractive index (Kerr-effect) of certain materials. The shorter, i.e. intenser, an incoming pulse is in the material, the stronger self-focusing it experiences. In combination with an aperture such a material ensures that short pulses will be amplified preferably. In combination with chirped mirrors that have been invented in 1990 (Szipocs et al. 1994) it was possible to demonstrate pulse durations shorter than 10 fs with this method. Chirped mirrors (see Figure 2, 3) are coated in a way that different wavelengths penetrated to a differing depth of the layers. The word “chirp” refers to the fact that frequencies - like in the singing of birds - are changing instantaneously. An incoming light pulse whose spectral components are asynchron (chirped) can be compressed to the minimum pulse duration which is supported by its spectrum. Using techniques of this kind, today it's possible to build oscillators whose optical spectrum comprises nearly a full octave, so that the high frequency part of the spectrum oscillates twice that fast as the low frequency one. In the example discussed here, the spectrum reaches from 450 to 1050 nm, supporting a pulse duration of about 5 fs. The ability to generate ultra-short laser pulses is not the only aspect that puts mode locked lasers in the focus of science. The lasers resonator emits pulses with a repetition rate corresponding to the inverse round trip time in the cavity. This repetition rate therefore depends on the length of the laser cavity and typical values are on the order of 100 MHz. It has to be emphasized that this periodicity in time translates to a periodicity in the Fourier domain: The spectrum of the pulse train emitted by a mode locked oscillator is not continuous but consists of perfectly equidistant spectral lines. This phenomenon was dubbed “frequency comb” and in 2005 Theodor Hänsch was awarded the Nobel Prize for the proposal to use this unique feature for frequency resolved spectroscopy (Reichert 1999) (Telle et al. 1999). Nowadays it is the basis for uncountable applications in the high resolution spectroscopy given that the spacing of the comb lines for a standard mode locked oscillator emitting at 800 nm with a repetition rate of 100 MHz is only 0.00021 nanometers.

Simultaneously the frequency comb found an important application also in the synthesis of ultrashort laser pulses themselves. Relying on the measurement of difference frequencies between the individual comb lines one can ensure that every laser pulse emitted has an identical evolution of the electric field underneath the intensity envelope, the so-called “waveform”. Figure 1 highlights the importance of this technology for ultrashort laser pulses containing only few optical cycles.

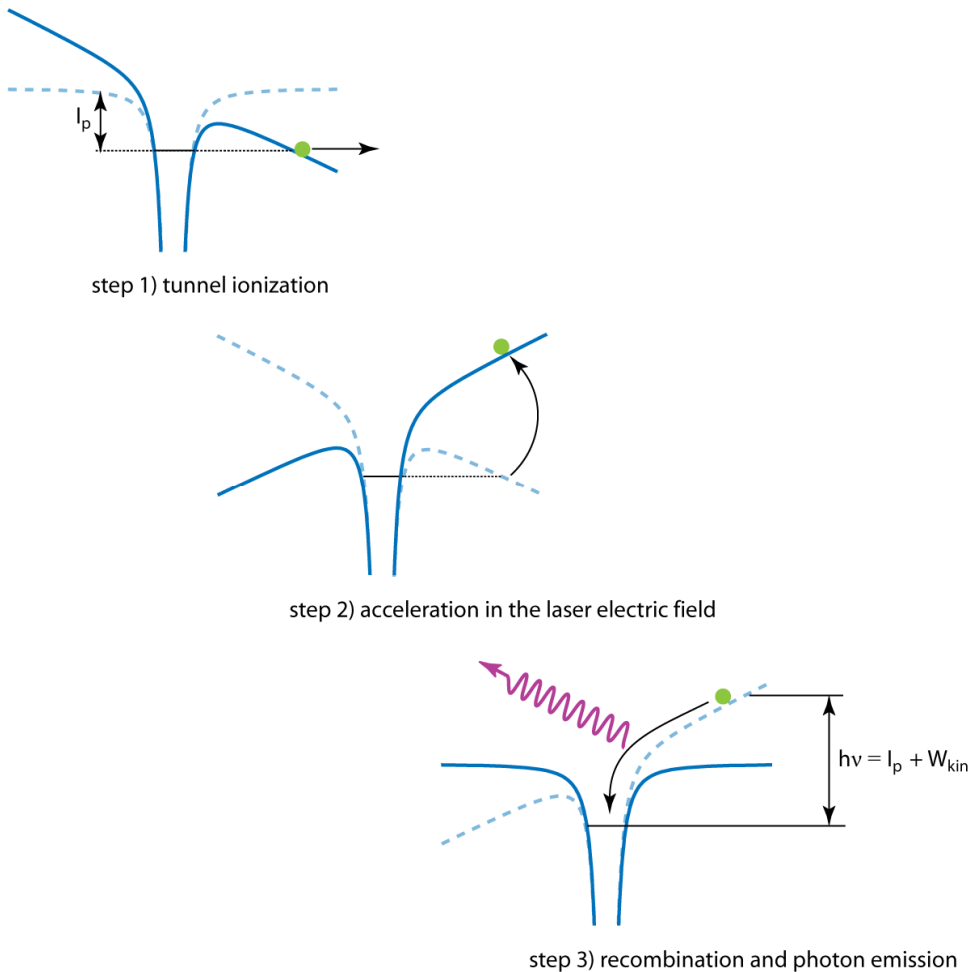


Fig. 1. High harmonic generation (HHG)

Ultra intense laser pulses can ionize matter but the released electrons not necessarily leave the atom forever. The "3-step-model" invented by Corkum in 1993 describes this process. If the field strength of the laser pulses is strong enough to suppress the coulomb potential binding the electron to the core, the electrons can tunnel out of the atom (step 1). This free electron is then accelerated in the laser electric field (step 2) that reverts its sign once per period and thus forces the electron to return to its parent ion with a non-vanishing probability and recombine there under the emission of radiation. This process is unlikely but not impossible; the conversion efficiencies reported are about  $10^{-7}$ . Still this technique offers amazing possibilities to the experimentalist due to the unique time structure of the emitted pulses (counted in attoseconds) and the high photon energies that can be achieved (up to keV photons are demonstrated).

Modern mode locked oscillators achieve Nano-Joule pulse energies which in turn can be focused to intensities in the range of  $\sim 10^{11}$  W/cm<sup>2</sup> insufficient to cause nonlinear response of even ionization of matter (effective ionization of rare gases sets in at laser intensities around  $10^{14}$  W/cm<sup>2</sup>).

Pushing ultrafast technology towards intensities above this threshold requires additional efforts (Figure 2). Usually mode locked oscillators are followed by amplification stages that work according to the „chirped-pulse-amplification“ scheme (STRICKLAND & MOUROU 1985). The individual spectral components are dispersed and thus the laser pulses are stretched in time leading to drastically reduced peak intensity while maintaining the coherence properties. Subsequently the pulses are amplified in a second laser active medium (crystal or fiber). For this purpose the repetition rate of the oscillator pulse train has to be reduced drastically since the amplifying medium has to fully absorb the thermal energy that goes along with the million fold amplification of the pulses. The difficult handling of the increasing average power and the lack of high-repetitive pump lasers currently render repetition rates higher than a few kilohertz impossible.

For selection of the individual pulses out of the high-repetitive oscillator pulse train contemporary laser systems rely on “pulse-picking” arrangements employing the electro optic effect in a Pockels-cell between orthogonal polarizing filters. After the amplification process (see Fig. 2) the timing of the individual wavelength components is rectified to achieve a short laser pulse by grating- or prism compressors frequently in combination with chirped mirrors.

Unfortunately the amplification process not only reduces the repetition rate but – due to the natural gain window of the laser active medium- also the spectral bandwidth of the laser pulses. To achieve near-single cycle pulses the amplification has to be accompanied by a system that coherently generates additional frequency components on either side of the spectrum. The commonly used systems to achieve that are based on a nonlinear process that was dubbed “self-phase-modulation”. The nonlinear interaction of the electric field of energetic laser pulses with rare gases can yield spectra spanning from 450 to 1100 nm covering the whole range of human vision (cp. Fig. 3). By the use of additional chirped mirrors to once again synchronize the individual spectral components state-of-the-art ultrashort pulse laser systems achieve mJ pulse energies within less than 3,5 femtoseconds (A. L. Cavalieri, E. Goulielmakis, et al. 2007) duration. Well focused, no matter can withstand the strength of the electric field that oscillates only 1,5 times during such a pulse reaching intensities up to  $10^{18}$  W/cm<sup>2</sup>.

If such pulses are used for the generation of high-order harmonics, as mentioned before (see also Fig. 1), one has to distinguish between pulses with many cycles of the driver wave and few-cycle pulses. In the first case, a train of attosecond pulses is generated, which requests that the sub-cycle electron-light interaction is repeated over several oscillation cycles under exactly the same conditions. Only then, no information is lost during the accumulation of attosecond signals from different cycles of the laser pulse, since the characteristics of the interaction products (e.g. the properties of photoelectrons) are the same. The same request has to be fulfilled when a train of attosecond pulses being separated by half-cycle of the driving laser field is used in a cross correlation with the driving field itself. Even if the laser field had constant amplitude over many cycles, metrology would be restricted to processes being shorter than one half-cycle of the laser field.

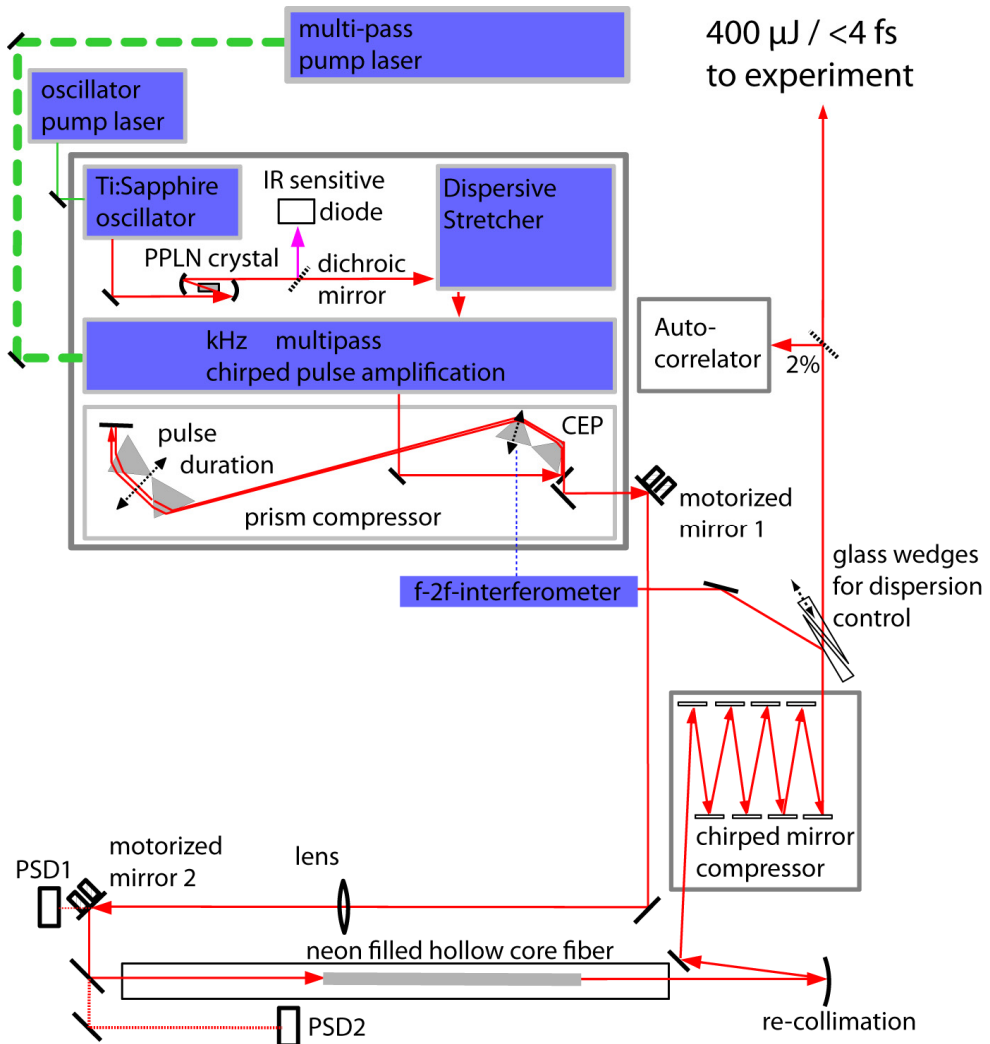


Fig. 2. Ultrafast Laser

A continuously pumped mode-locked oscillator emits broadband pulses of few-femtosecond duration. These pulses are dispersed in time up to ps duration to reduce their peak-intensity and subsequently amplified in a multi-pass amplifier pumped by a multi-mJ Q-switch laser usually operating with kHz repetition rates. Later a prism-compressor rectifies the timing of the different wavelength components reducing the pulse duration again to 10s of femtoseconds. The nonlinear process of self-phase-modulation that takes place in a gas filled hollow-core fiber adds additional spectral components on either side of the amplified spectrum and a set of chirped mirrors converts the octave spanning output spectrum (cp. Figure 3) into ultrashort visible light pulses.

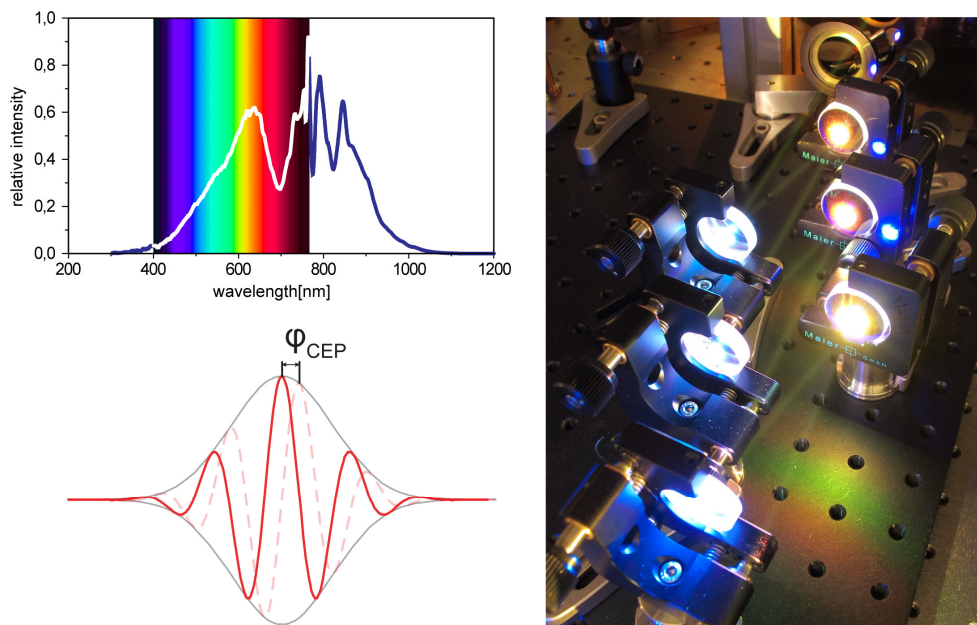


Fig. 3. Ultrashort Laser pulses

As a consequence of the uncertainty relation short laser pulses require an extended spectrum. To synthesize laser pulses consisting of little more than one cycle of the electric field the spectrum has to span an octave (it has to coherently contain frequencies between  $\omega$  and  $2\omega$ ). Given the human eyes sensitivity being restricted to a range between 400 nm and 750 nm, single cycle laser pulses appear white as seen on the photograph that shows the scattered light if such pulses are reflected off the multilayer mirrors mentioned in the text. The graph compares the measured spectrum of such pulses with the sensitivity range of human sight (rainbow shaded area). The plot below highlights how for such short pulses the phase (carrier-envelope-phase CEP) between the intensity envelope function and the oscillations of the carrier wave defines the waveform and thus the action of the pulse in nonlinear processes.

Multi-cycle attosecond metrology (Ferenc Krausz 2009) can be used to investigate periodically-repeated recollisions which are driven by the multi-cycle light wave and trigger the process under study in a correlated manner in the same interaction within each wave cycle. One of the two correlated processes can serve as a clock for the other. Inventing this concept, (Niikura et al. 2002) demonstrated that the vibrational motion of a molecule triggered by ionization clocks the recollision electron and (Niikura et al. 2003) showed the case vice versa for a diatomic molecule.

In most laser systems, multi-cycle pulses have a Gaussian envelope which results in attosecond pulses with varying intensity and photon energy within the pulse train generated, If these pulses trigger even simple processes a retrieval of the process gets highly complicated.

Therefore, employing few-cycle pulses for the synthesis of high-order-harmonics (Fig. 1) becomes highly significant. Given a smart choice of the phase between the intensity envelope and the oscillations of the carrier wave, waveforms can be generated where the intensity contrast between adjacent maxima of the electric field is maximized.

Since every single half cycle contributes to the emission of high-harmonics the highest intense half cycle apparently leads to the emission of the highest energetic photons. Filtering only this radiation out of the emitted spectrum thus yields coherent X-ray pulses with a duration corresponding to only a fraction of the driving lasers half-cycle duration. The shortest light bursts demonstrated ever have been generated in this manner: Lasting only about 80 attoseconds they contain roughly  $10^8$  photons per pulse at a central wavelength of 12 nm deep in the ultraviolet.

As has been discussed, the parameters of the driving laser pulse are extremely important for the characteristics of the generated harmonic radiation. Pulse duration, pulse energy, photon energy, and the number of attosecond bursts (single pulses or a train) are given by the properties of the driving laser pulse. In terms of the photon energy, the wavelength of the driving pulse is crucial. It may be counterintuitive on first glance but longer wavelength driver pulses are able to generate harmonics at higher photon energy (Vladislav S. Yakovlev et al. 2007) since the electron can accumulate more energy during its trip to the continuum (in the semi classical picture). Therefore the development of few-cycle laser pulses in the infrared spectral range is pursued by several groups in the world.

(Fuji et al. 2006) have demonstrated an OPA system whose pulses comprise only a few oscillation cycles at 2.1  $\mu\text{m}$  carrier wavelength. IR-driven HHG was pioneered by L'Huillier and co-workers (L'Huillier & Ph. Balcou 1993) and DiMauro and co-workers (Sheehy et al. 1999) followed by studies of (Bellini 2000), (Shan & Chang 2001). As pointed out by DiMauro, the favourable scaling of the maximum photon energy with driver wavelength may open the way to generating coherent light and attosecond pulses at photon energies substantially beyond the kiloelectronvolt frontier, which was reached recently with NIR few-cycle light (J. Seres et al. 2005). This follows from the fact that the cutoff, i.e. the highest achievable energy in HHG (see Fig 1) scales with the ponderomotive potential - the cycle averaged quiver energy an electron gains in an oscillating field - and therefore with the square of the driving wavelength.

Another technique that aims at increasing the flux of HHG especially at high photon energies is the so-called quasi phase-matching (QPM). In this approach one tries to overcome a problem that limits the conversion efficiency between driver and harmonic radiation. Dispersion in the non-linear medium used for HHG results in a phase mismatch between the driver wave in the visible and NIR spectral range and the short wavelength product of the conversion process. This leads to a destructive interference after a certain distance and reduces the flux of the harmonic radiation. Methods like using a modulated fiber (Christov et al. 1998) as a cell for the conversion process or using several targets to modulate the density of the conversion medium aim at bringing the both waves involved in cadence again. Especially for photon energies in the keV range, where the phase mismatch plays a crucial role, this method is of great importance.

### 3. Experiments and applications

If solids are used as a medium for the frequency upconversion to the XUV or x-ray regime, a different mechanism comes into play (Naumova et al. 2004); (Tsakiris et al. 2006); (Baeva et al. 2006)). At ultra-high intensities, the leading edge of a laser pulse can turn the solid surface into a high-density plasma where the electrons execute oscillations at the plasma-vacuum interface driven by the ultra-intense field. The light pulse bounces back from the surface as if it were reflected by an oscillating mirror with relativistic speed (Bulanov et al. 1994) (Lichters et al. 1996); (Gordienko et al. 2004); (Tsakiris et al. 2006). Since the time it takes the mirror to sweep through a half period of the reflected field is contracted when the mirror and the incident wave counter-propagate, the reflected wave is blue-shifted. The highest occurring frequency is then given by the maximum Doppler upshift of  $\Omega_{\max} - 4\gamma_{\max}^2\omega_L = 4(1 + a_0^2)\omega_L$ , with  $\gamma_{\max}$  being the relativistic factor and  $\omega_L$  and  $a_0$  being the frequency and vector potential of the laser pulse. The emission of high-frequency photons is confined to small fractions of the laser period resulting in a periodic train of high-energy photon bursts. In a few-cycle laser field the emission of the high-energy photons is confined to the central wave cycle. Spectral filtering may therefore lead to single attosecond pulse generation (Ferenc Krausz 2009).

After its advent at the beginning of the century attosecond science soon was recognized to have the potential to add unprecedented temporal resolution to most of the spectroscopic techniques that have been developed at large scale facilities and laboratories around the world. Particularly (photo-) electron spectroscopy (see Fig. 4) was enriched by the attosecond toolbox when *attosecond streaking spectroscopy* emerged (cp. Fig 5) (Hentschel et al. 2001) (Itatani et al. 2002) (R Kienberger et al. 2002) (Markus Kitzler et al. 2002) (R. Kienberger et al. 2004). Attosecond XUV pulses are focused onto the sample atoms (usually in the gas phase but also surface emission experiments have been demonstrated) causing the release of electrons with a kinetic energy corresponding to the energy difference between the XUV photon energy and the electron affinity. This kinetic energy is detected by measuring the electron arrival time on a detector placed in a known distance of the interaction volume as in common electron time of flight spectrometry. For this type of spectroscopy (in contrast to techniques involving trains of attosecond pulses) it is a prerequisite that the XUV pulse duration is significantly shorter than a quarter of the laser period. For lasers with a central wavelength on the red side of the visible spectrum that amounts to round 500 attoseconds and as a direct consequence of the uncertainty principle this goes along with a minimal spectral width of 3 eV at a central photon energy around 100 eV. Together with the spectral resolution around 1% that is typically achieved in time of flight spectrometers this sets a lower limit to the spectral distance of features that shall be resolved and contrasts to the spectral resolution in the meV range that e.g. synchrotron sources achieve. Instead attosecond science exhibits its full strength in resolving ultrafast processes: For the sake of three orders of magnitude in spectral resolution it adds about 12 orders of magnitude in temporal resolution compared to the most contemporary XUV technologies known from synchrotrons and linear accelerators.

If the photoelectrons are set free in the presence of the electric field of a laser pulse their final kinetic energy can be altered (Hentschel et al. 2001) (Itatani et al. 2002). The net momentum change imparted on the electrons depends on the time delay between the attosecond XUV and the laser pulse (M. Schultze et al. 2011). A number of time of flight spectra is recorded with incrementally increased delay between the two pulses around the temporal coincidence.



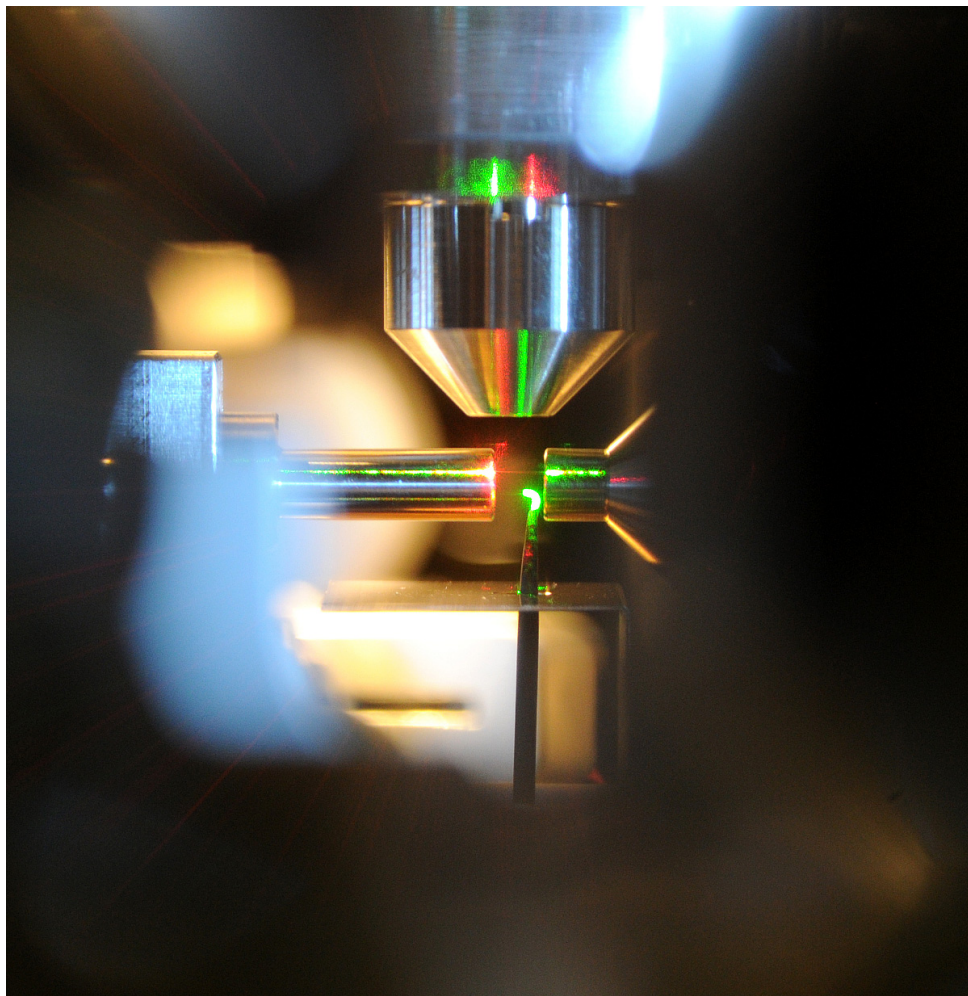
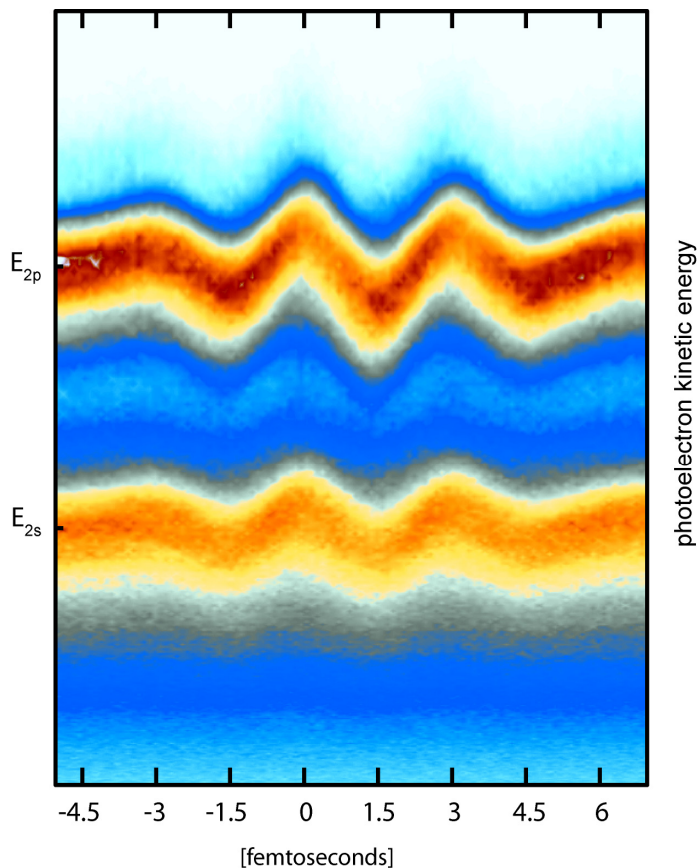


Fig. 4. Attosecond Spectroscopy

Experiments designed to make use of the unprecedented temporal resolution that ultrashort laser pulses can offer usually follow the „pump-probe-scheme“. A first laser pulse triggers the process under scrutiny and a second pulse probes the evolution after a defined time interval. From a sequence of a number of such measurements with variable temporal delay between the two pulses one can reconstruct the temporal evolution. The picture shows the inside of a vacuum chamber where such pump-probe experiments are performed based on the combination of visible laser pulses as pump- and attosecond XUV pulses as probe events. It is used to e.g. explore the electron dynamics in the atomic core. The gaseous sample streams out of the black nozzle into the interaction region and the released electrons are collected by a time-of-flight detector (cone from above) while the ionic fragments are detected by a mass spectrometer along the horizontal axis. The image was taken on the beam axis of the collinear propagating laser and xuv pulses .



Delay between "pump" and "probe" pulse

Fig. 5. Attosecond Streaking

If an attosecond XUV pulse releases photoelectrons from an atom while at the same time a strong visible laser pulse shines on the sample a streaking spectrogram can be recorded. When leaving the atom, the photoelectrons have a characteristic kinetic energy that corresponds to the difference between the electron affinity and the XUV photon energy. In the presence of an ultrastrong laser pulse this kinetic energy is modified by the electric field (as an electron is accelerated towards the positively charged plate in a capacitor). Since the electric field oscillates the change in kinetic energy depends on the difference in arrival time of the two pulses involved. The figure shows a measurement where the photoelectrons kinetic energy spectrum (along the vertical axis) is recorded for about 100 different delay settings between the two pulses (along the horizontal axis, the delay is given in femtoseconds). The result is an "oscilloscope" recording of the electric field of the laser pulse. Here Neon gas was ionized by an attosecond XUV pulse with photon energy about 120 eV (corresponding to 10 nm wavelength) and a duration of 200 attoseconds. The photon energy is sufficient to release electrons from either the  $2p$  or the  $2s$  shell of Neon thus the measurement allows to investigate potential timing differences in the electron emission (M Schultze et al. 2010).

With a step size roughly matching the XUV pulse duration the resulting spectrogram can look as in Fig. 5 and it is immediately evident that this so called “streaking spectroscopy” reveals the temporal evolution of the laser electric field (more precisely its vector potential) (E Goulielmakis et al. 2004) (Fiess et al. 2010). Beyond that, this sort of spectrograms resembles the signal of frequency resolved optical gating devices (FROG) known from laser science (Trebino & Kane 1993); (Mairesse & Quéré 2005); (Justin Gagnon & Vladislav S. Yakovlev 2009). Since their origin is a cross-correlation of two unknown pulses recorded with a known response function (here the photoionization) adaptive algorithms can extract all relevant parameters of **both** pulses involved (Michael Hofstetter et al. 2011). Such algorithms identify the temporal structure of the attosecond pulses and their temporal and spectral phase and thus fully characterize these light bursts that are the shortest signals that can be synthesized in the laboratory. It is an interesting feature of this streaking concept that the reconstruction actually explores the temporal structure of the released photoelectron wavepacket rather than of the light pulse and thus gives access to characteristics of the photo emission process that so far remained inaccessible to experimental physics. For photon energies that are sufficiently high to address two or more individual atomic orbitals and thus liberate electrons from different energetic states (as shown in Fig. 5) the streaking technique facilitates the investigation of subtle timing differences between them. The growing ability to shape the XUV emission according to the requirements of a particular experiment extends the technique to a wider spectral range. With specially tailored attosecond XUV pulses it was possible to discover that the emission of electrons from the neon 2s shell precedes the liberation of electrons originally bound in the 2p shell by about 20 attoseconds, the shortest time interval ever sampled directly (M Schultze et al. 2010).

#### 4. Conclusion

Attosecond pulses in general allow taking snapshots of the fastest processes known in the microcosm. They are permitted to resolve processes, so far only predicted, in real time and facilitated the discovery of ultra-fast mechanisms that are not yet fully understood regarding their physical origin.

The quantum mechanical tunnel effect belongs to the first category. As an implementation of *attosecond tunneling spectroscopy* a first experiment demonstrated how adding attosecond tools to ion mass spectrometry can shine -time-resolved- light on very fundamental processes. According to the explanation of Leonid Keldysh the probability for tunnel ionization is high at the maxima of the laser electric field and thus leads to a stepwise evolution of the number of ionized atoms exposed to intense light (Mishima et al. 2002) at the same time the time structure of subsequent Auger decay cascades in Xenon and Krypton have been resolved (Uiberacker et al. 2007). So far only known from (indirect) line width measurements, attosecond pulses turned out to be the only tool that allows unraveling complex and intertwined processes in real time.

The investigation of transport effects of electrons in a solid is an example for the second category (A. L. Cavalieri, Müller, et al. 2007). A timing difference of electrons reaching the surface after photoionization was discovered that depends on the initial state of the released electron. It was found that electrons originating from the 4f band of a tungsten crystal take about 50 attoseconds longer to reach the surface compared to those starting in the conduction band. Is that solely a transport effect or signature of the crystal band structure?

Recently attosecond science was successfully combined with *transient absorption spectroscopy*. This experimental approach, that revolutionized quantum chemistry after the advent of pulsed laser sources, was used to track the oscillations of a coherent superposition in the valence shell of atoms in real time (Eleftherios Goulielmakis et al. 2010).

## 5. References

- Anon, 1994. *Femtochemistry: ultrafast dynamics of the chemical bond*, Available at: [http://books.google.com/books?hl=de&lr=&id=T8IZ1aa\\_FRkC&pgis=1](http://books.google.com/books?hl=de&lr=&id=T8IZ1aa_FRkC&pgis=1) [Accessed October 14, 2011].
- Baeva, T., Gordienko, S. & Pukhov, A., 2006. Relativistic plasma control for single attosecond x-ray burst generation. *Physical Review E*, 74(6). Available at: <http://pre.aps.org/abstract/PRE/v74/i6/e065401> [Accessed October 6, 2011].
- Bellini, M., 2000. Generation of widely tunable harmonic pulses in the UV and VUV from a NIR optical parametric amplifier. *Applied Physics B Lasers and Optics*, 70(6), pp.773-776. Available at: <http://www.mendeley.com/research/generation-widely-tunable-harmonic-pulses-uv-vuv-nir-optical-parametric-amplifier/> [Accessed October 14, 2011].
- Bulanov, S.V., Naumova, N.M. & Pegoraro, F., 1994. Interaction of an ultrashort, relativistically strong laser pulse with an overdense plasma. *Physics of Plasmas*, 1(3), p.745. Available at: <http://link.aip.org/link/?PHPAEN/1/745/1> [Accessed October 14, 2011].
- Cavalieri, A. L., Müller, N., et al., 2007. Attosecond spectroscopy in condensed matter. *Nature*, 449(7165), pp.1029-32. Available at: <http://dx.doi.org/10.1038/nature06229> [Accessed July 19, 2011].
- Cavalieri, A. L., Goulielmakis, E., et al., 2007. Intense 1.5-cycle near infrared laser waveforms and their use for the generation of ultra-broadband soft-x-ray harmonic continua. *New Journal of Physics*, 9(7), pp.242-242. Available at: <http://stacks.iop.org/1367-2630/9/i=7/a=242?key=crossref.c8853a078987c15542c906c0692977b1> [Accessed June 29, 2011].
- Christov, I., Kapteyn, H. & Murnane, M., 1998. Dispersion-controlled hollow core fiber for phase matched harmonic generation. *Optics Express*, 3(10), p.360. Available at: <http://www.opticsexpress.org/abstract.cfm?URI=oe-3-10-360> [Accessed October 14, 2011].
- Christov, I., Murnane, M. & Kapteyn, H., 1997. High-Harmonic Generation of Attosecond Pulses in the "Single-Cycle" Regime. *Physical Review Letters*, 78(7), pp.1251-1254. Available at: [http://prl.aps.org/abstract/PRL/v78/i7/p1251\\_1](http://prl.aps.org/abstract/PRL/v78/i7/p1251_1) [Accessed August 3, 2011].
- Ferray, M. et al., 1988. Multiple-harmonic conversion of 1064 nm radiation in rare gases. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 21(3), p.L31-L35. Available at: <http://stacks.iop.org/0953-4075/21/i=3/a=001> [Accessed July 27, 2011].
- Fiess, M. et al., 2010. Versatile apparatus for attosecond metrology and spectroscopy. *The Review of scientific instruments*, 81(9), p.093103. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20886972> [Accessed September 19, 2011].
- Fuji, T. et al., 2006. Parametric amplification of few-cycle carrier-envelope phase-stable pulses at 2.1  $\mu\text{m}$ . *Optics Letters*, 31(8), p.1103. Available at: <http://ol.osa.org/abstract.cfm?URI=ol-31-8-1103> [Accessed October 14, 2011].

- Gagnon, Justin & Yakovlev, Vladislav S., 2009. The robustness of attosecond streaking measurements. *Optics Express*, 17(20), p.17678. Available at: <http://www.opticsexpress.org/abstract.cfm?URI=oe-17-20-17678> [Accessed October 7, 2011].
- Gordienko, S. et al., 2004. Relativistic Doppler Effect: Universal Spectra and Zeptosecond Pulses. *Physical Review Letters*, 93(11). Available at: <http://prl.aps.org/abstract/PRL/v93/i11/e115002> [Accessed July 21, 2011].
- Goulielmakis, E et al., 2004. Direct measurement of light waves. *Science (New York, N.Y.)*, 305(5688), pp.1267-9. Available at: <http://www.sciencemag.org/content/305/5688/1267.abstract> [Accessed October 6, 2011].
- Goulielmakis, Eleftherios et al., 2010. Real-time observation of valence electron motion. *Nature*, 466(7307), pp.739-43. Available at: <http://dx.doi.org/10.1038/nature09212> [Accessed July 18, 2011].
- Hentschel, M. et al., 2001. Attosecond metrology. *Nature*, 414(6863), pp.509-13. Available at: <http://dx.doi.org/10.1038/35107000> [Accessed October 7, 2011].
- Hofstetter, Michael et al., 2011. Attosecond dispersion control by extreme ultraviolet multilayer mirrors. *Optics Express*, 19(3), p.1767. Available at: <http://www.opticsexpress.org/abstract.cfm?URI=oe-19-3-1767> [Accessed October 14, 2011].
- Itatani, J. et al., 2002. Attosecond Streak Camera. *Physical Review Letters*, 88(17). Available at: <http://prl.aps.org/abstract/PRL/v88/i17/e173903> [Accessed August 7, 2011].
- Kienberger, R. et al., 2004. Atomic transient recorder. *Nature*, 427(6977), pp.817-21. Available at: <http://dx.doi.org/10.1038/nature02277> [Accessed August 11, 2011].
- Kienberger, R et al., 2002. Steering attosecond electron wave packets with light. *Science (New York, N.Y.)*, 297(5584), pp.1144-8. Available at: <http://www.sciencemag.org/content/297/5584/1144.abstract> [Accessed July 18, 2011].
- Kitzler, Markus et al., 2002. Quantum Theory of Attosecond XUV Pulse Measurement by Laser Dressed Photoionization. *Physical Review Letters*, 88(17). Available at: <http://prl.aps.org/abstract/PRL/v88/i17/e173904> [Accessed August 7, 2011].
- Krausz, Ferenc, 2009. Attosecond physics. *Reviews of Modern Physics*, 81(1), pp.163-234. Available at: [http://rmp.aps.org/abstract/RMP/v81/i1/p163\\_1](http://rmp.aps.org/abstract/RMP/v81/i1/p163_1) [Accessed July 10, 2011].
- Lichters, R., Meyer-ter-Vehn, J. & Pukhov, A., 1996. Short-pulse laser harmonics from oscillating plasma surfaces driven at relativistic intensity. *Physics of Plasmas*, 3(9), p.3425. Available at: <http://link.aip.org/link/?PHPAEN/3/3425/1> [Accessed October 14, 2011].
- L'Huillier, A. & Balcou, Ph., 1993. High-order harmonic generation in rare gases with a 1-ps 1053-nm laser. *Physical Review Letters*, 70(6), pp.774-777. Available at: [http://prl.aps.org/abstract/PRL/v70/i6/p774\\_1](http://prl.aps.org/abstract/PRL/v70/i6/p774_1) [Accessed October 14, 2011].
- Mairesse, Y. & Quéré, F., 2005. Frequency-resolved optical gating for complete reconstruction of attosecond bursts. *Physical Review A*, 71(1). Available at: <http://pra.aps.org/abstract/PRA/v71/i1/e011401> [Accessed August 7, 2011].

- McPherson, A. et al., 1987. Studies of multiphoton production of vacuum-ultraviolet radiation in the rare gases. *Journal of the Optical Society of America B*, 4(4), p.595. Available at: <http://josab.osa.org/abstract.cfm?URI=josab-4-4-595> [Accessed July 27, 2011].
- Mishima, K. et al., 2002. Generalization of Keldysh's theory. *Physical Review A*, 66(3). Available at: <http://pra.aps.org/abstract/PRA/v66/i3/e033401> [Accessed October 14, 2011].
- Naumova, N.M. et al., 2004. Relativistic Generation of Isolated Attosecond Pulses in a  $\lambda^{\{3\}}$  Focal Volume. *Physical Review Letters*, 92(6). Available at: <http://prl.aps.org/abstract/PRL/v92/i6/e063902> [Accessed October 14, 2011].
- Niikura, H. et al., 2002. Sub-laser-cycle electron pulses for probing molecular dynamics. *Nature*, 417(6892), pp.917-22. Available at: <http://dx.doi.org/10.1038/nature00787> [Accessed June 13, 2011].
- Niikura, H. et al., 2003. Probing molecular dynamics with attosecond resolution using correlated wave packet pairs. *Nature*, 421(6925), pp.826-9. Available at: <http://dx.doi.org/10.1038/nature01430> [Accessed July 23, 2011].
- Paul, P.M. et al., 2001. Observation of a train of attosecond pulses from high harmonic generation. *Science (New York, N.Y.)*, 292(5522), pp.1689-92. Available at: <http://www.sciencemag.org/content/292/5522/1689.abstract> [Accessed August 7, 2011].
- Reichert, J., 1999. Measuring the frequency of light with mode-locked lasers. *Optics Communications*, 172(1-6), pp.59-68. Available at: [http://dx.doi.org/10.1016/S0030-4018\(99\)00491-5](http://dx.doi.org/10.1016/S0030-4018(99)00491-5) [Accessed October 14, 2011].
- Strickland, D. & Mourou, G., 1985. Compression of amplified chirped optical pulses☆. *Optics Communications*, 56(3), pp.219-221. Available at: [http://dx.doi.org/10.1016/0030-4018\(85\)90120-8](http://dx.doi.org/10.1016/0030-4018(85)90120-8) [Accessed June 20, 2011].
- Schultze, M et al., 2010. Delay in photoemission. *Science (New York, N.Y.)*, 328(5986), pp.1658-62.
- Schultze, M. et al., 2011. State-of-the-art attosecond metrology. *Journal of Electron Spectroscopy and Related Phenomena*, 184(3-6), pp.68-77. Available at: <http://dx.doi.org/10.1016/j.elspec.2011.01.003> [Accessed July 18, 2011].
- Seres, J. et al., 2005. Laser technology: source of coherent kiloelectronvolt X-rays. *Nature*, 433(7026), p.596. Available at: <http://dx.doi.org/10.1038/433596a> [Accessed August 3, 2011].
- Shan, B. & Chang, Z., 2001. Dramatic extension of the high-order harmonic cutoff by using a long-wavelength driving field. *Physical Review A*, 65(1). Available at: <http://pra.aps.org/abstract/PRA/v65/i1/e011804> [Accessed July 4, 2011].
- Sheehy, B. et al., 1999. High Harmonic Generation at Long Wavelengths. *Physical Review Letters*, 83(25), pp.5270-5273. Available at: [http://prl.aps.org/abstract/PRL/v83/i25/p5270\\_1](http://prl.aps.org/abstract/PRL/v83/i25/p5270_1) [Accessed July 27, 2011].
- Szipocs, R. et al., 1994. Chirped multilayer coatings for broadband dispersion control in femtosecond lasers. *Optics Letters*, 19(3), p.201. Available at: <http://ol.osa.org/abstract.cfm?URI=ol-19-3-201> [Accessed September 16, 2011].

- Telle, H.R. et al., 1999. Carrier-envelope offset phase control: A novel concept for absolute optical frequency measurement and ultrashort pulse generation. *Applied Physics B: Lasers and Optics*, 69(4), pp.327-332. Available at: <http://www.springerlink.com/content/101jv0l4xejwbw6j/> [Accessed October 14, 2011].
- Trebino, R. & Kane, D.J., 1993. Using phase retrieval to measure the intensity and phase of ultrashort pulses: frequency-resolved optical gating. *Journal of the Optical Society of America A*, 10(5), p.1101. Available at: <http://josaa.osa.org/abstract.cfm?URI=josaa-10-5-1101> [Accessed October 14, 2011].
- Tsakiris, G.D. et al., 2006. Route to intense single attosecond pulses. *New Journal of Physics*, 8(1), pp.19-19. Available at: <http://stacks.iop.org/1367-2630/8/i=1/a=019> [Accessed October 14, 2011].
- Uiberacker, M. et al., 2007. Attosecond real-time observation of electron tunnelling in atoms. *Nature*, 446(7136), pp.627-32. Available at: <http://dx.doi.org/10.1038/nature05648> [Accessed June 24, 2011].
- Yakovlev, Vladislav S., Ivanov, Misha & Krausz, Ferenc, 2007. Enhanced phase-matching for generation of soft X-ray harmonics and attosecond pulses in atomic gases. *Optics Express*, 15(23), p.15351. Available at: <http://www.opticsexpress.org/abstract.cfm?URI=oe-15-23-15351> [Accessed October 14, 2011].

# Measurement of Ultrashort Optical Pulses

Yuqiang Deng<sup>1,\*</sup>, Qing Sun<sup>1</sup>, Shiyong Cao<sup>2</sup>,  
Jing Yu<sup>1</sup>, Ching-yue Wang<sup>3</sup> and Zhigang Zhang<sup>4</sup>

<sup>1</sup>*Optics Division, National Institute of Metrology, Beijing,*

<sup>2</sup>*Time and Frequency Center, National Institute of Metrology, Beijing,*

<sup>3</sup>*School of Precision Instrument and Optoelectronics Engineering,*

*Tianjin University, Tianjin,*

<sup>4</sup>*School of Electronics Engineering and Computer Science, Peking University, Beijing,*

*P. R. China*

## 1. Introduction

An ultrashort optical pulse has an extremely short duration period, an extremely broad spectral bandwidth, and an extremely high peak power [1]; it has therefore been widely used in a variety of applications, such as ultrafast pumping and detection, time-resolved spectroscopy, optical telecommunications, ultra-fine microfabrication, nonlinear optics and femtosecond chemistry. Femtosecond optical pulses have also brought revolutions in contemporary metrology [2], including time and frequency standards [3], terahertz metrology [4], and ultrafast electric pulses characterization [5]. Time domain waveform and pulse width are key parameters of the ultrashort optical pulses, because they directly affect experimental results acquired by use of ultrashort optical pulses. Experimental data cannot be deemed credible unless the waveform and pulse width are determined. Accurate knowledge of the temporal shape of optical pulses is therefore crucial to scientific research.

In the past three decades, several measurement techniques for ultrashort optical pulse have been developed. Autocorrelation (AC) [6], frequency-resolved optical gating (FROG) [7] and spectral phase interferometry for direct electric-field reconstruction (SPIDER) [8] are most commonly used. Autocorrelation is simple and convenient, but can give only the autocorrelation width; the waveform and phase is difficult to obtain. FROG is a two-dimensional measurement technique, pulse waveform and phase can be retrieved from the FROG trace, but an iterative procedure is needed. SPIDER can directly extract spectral phase and reconstruct pulse waveform; it is therefore suitable for accurate and fast measurement of ultrashort optical pulses, especially for the optical pulses in femtosecond range. In this chapter, we would introduce our technique for spectral phase retrieval to improve the accuracy of femtosecond optical pulses measurement, and demonstrate a versatile autocorrelator for ultrashort optical measurement, with which the time domain measurement range is greatly broaden.

---

\* Corresponding Author



## 2. Characterization of femtosecond pulses with wavelet-transform for spectral phase retrieval

The spectrum of a pulse can easily be measured with a spectrometer. The pulse would be completely known, if we could, in addition, determine the phase across the spectrum [9]. SPIDER is a technique for measurement spectral phase of femtosecond optical pulses. In SPIDER setup, two replicas of the input pulse to be characterized are generated with a fixed time delay between them. These two replica pulses are then upconverted by sum-frequency mixing with a strongly chirped pulse derived from the same original input pulse. Because the two replica pulses are separated in time domain, they interact with different parts of the chirped pulse and are therefore upconverted to different frequencies. From the interferogram of these spectral shearing pulses, it is possible to extract the amplitude and phase of the initial pulse by using an algebraic inversion algorithm [2].

### 2.1 Optical schematics of SPIDER

Our home made SPIDER setup is shown in Fig. 1. Figure 1(a) is the schematic optical path. An input femtosecond laser pulse passes through an aperture A1, and then is reflected by a reflective mirror M1. The reflected pulse is split into two pulses by a beam splitter BS1. One of the split pulse is stretched into a strongly chirped pulse, and another is fed into a Michelson-type interferometer and then replicate into two replicas. The Michelson-type interferometer consists of two beam splitters BS2, BS3 and two corner mirrors TS2, TS3. A fixed time delay between the replicas is generated by tuning the translation stage TS2. The strongly chirped pulse and the replicas are focused by a paraboloid mirror PM, and then sum-frequency generation occurs in a sum-frequency crystal CR. A translation stage TS1 is used for tuning the relative time delay between the strongly chirped pulse and the two replicas. The two replica pulses interact with different parts of the chirped pulse and are upconverted to two different frequencies. The spectral interferogram of the frequency upconverted pulses is recorded by a fiber spectrometer (HR 4000, Ocean Optics Inc.). The spectral phase of the input pulse can be retrieved from the measured spectral interferogram. Our home-made experimental optical setup is shown in Fig. 1(b), with the optical path superposed.

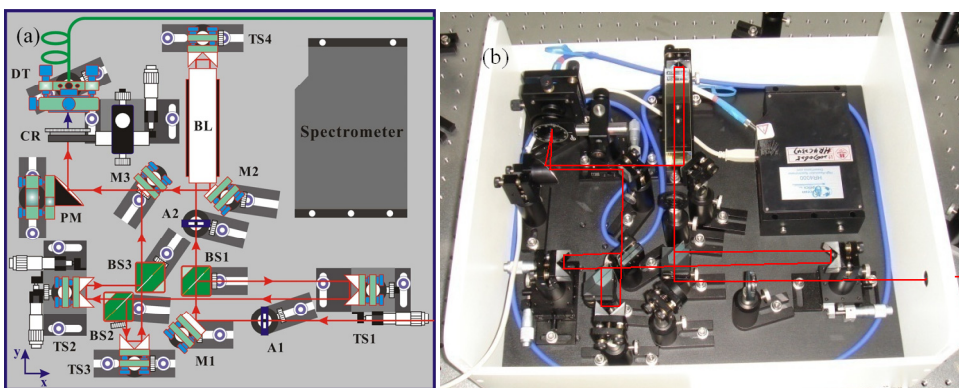


Fig. 1. Our home-made SPIDER setup and the optical path. (a) Optical schematic, (b) experimental setup.

## 2.2 Spectral phase retrieval with a wavelet-transform

In the traditional phase retrieval algorithm, spectral phase is extracted from the filtered alternating current component of the Fourier transform [8]. The filter is set by manual selection and adjustment. A problem is that different widths or shapes of the filter produce different spectral phases [10]. Thus the uncertainty of the spectral phase contributes to the uncertainty of the reconstructed pulses.

To reduce the uncertainty component coming from the filter with Fourier transform technique, we introduce a wavelet-transform for spectral phase retrieval of femtosecond optical pulses [11]. The phase is directly extracted from the ridge of wavelet-transform. There are no filters in this procedure so that the uncertainty from the filter is eliminated. In what follows a demonstration of this procedure will be shown.

We have measured an ultrashort optical pulse train emitted from a Ti:sapphire laser (Micra-5, Coherent Inc.). The average output power of the laser is 360 mW after a pulse compressor. The repeat frequency is 82 MHz, and the central wavelength is 800 nm with a spectral bandwidth (FWHM) of 100 nm. We perform a SPIDER measurement with our home-made SPIDER setup. The measured spectral interferogram is shown in Fig. 2.

Wavelet-transform was applied on the measured spectral interferogram, and the time and frequency distributions are exhibited on a two-dimensional plane. The intensity map and phase map are shown in Figs. 3(a) and 3(b), respectively.

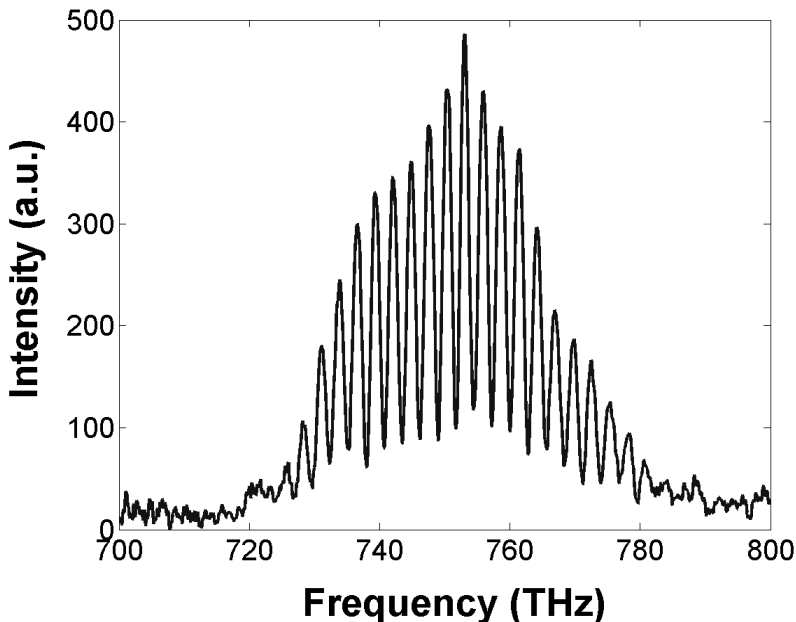


Fig. 2. Measured spectral interferogram.

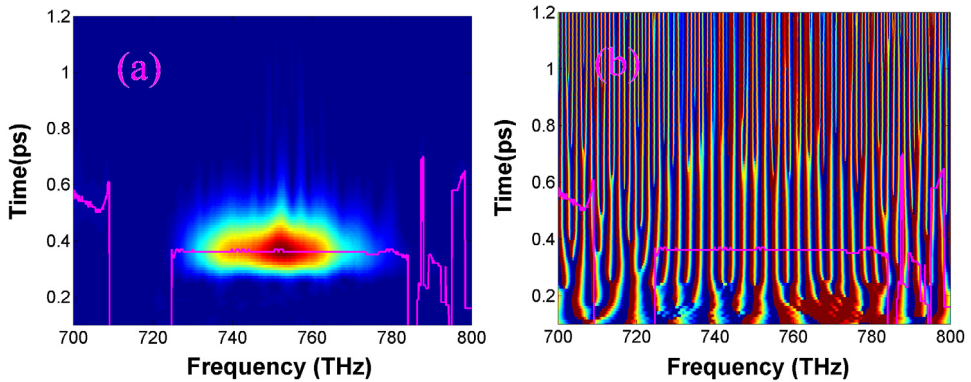


Fig. 3. Wavelet-transform of spectral interferogram. (a) Intensity topography, (b) phase topography. The ridge of wavelet-transform is indicated with a pink coloured line.

We search for the maximum values from the intensity topography along each frequency column. Connecting the positions of the maximum values at each frequency point constructs the ridge line of the wavelet-transform, which is superposed on Fig. 3(a) with a pink coloured line. Then we project the position of ridge from intensity topography (Fig. 3(a)) on the phase topography, as is shown in Fig. 3(b).

The phase of the spectral interferogram was directly extracted from the phase topography at the position of the ridge. With the extracted interferometry phase, the spectral phase was obtained with a concatenation algorithm [8], as is shown in Fig. 4.

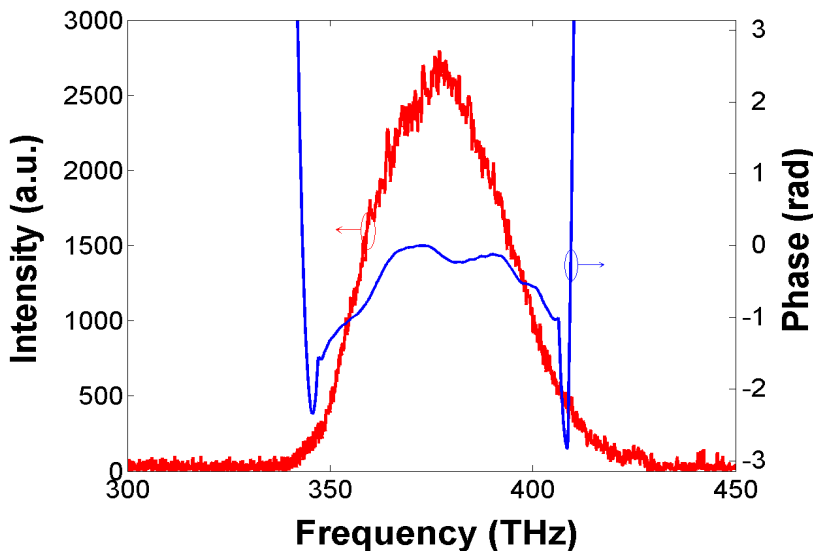


Fig. 4. Measured spectrum and retrieved spectral phase.

### 2.3 Pulse waveform reconstruction and autocorrelation simulation

Figure 4 shows the measured spectrum with a spectrometer (HR4000 CG-UV-NIR, Ocean Optics Inc.) and the retrieved spectral phase with wavelet-transform technique. The electric field (Fig. 5(a)) and waveform (Fig. 5(b)) of the femtosecond optical pulse are reconstructed from the spectrum and spectral phase with an inverse Fourier transform technique. The pulse width (FWHM) shown in Fig. 5(b) is 18.2 fs.

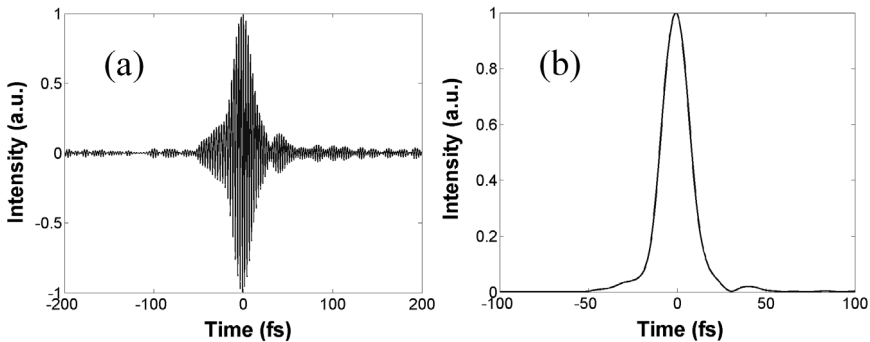


Fig. 5. Reconstructed electric field and waveform. (a) Electric field, (b) waveform.

From the reconstructed electric field in Fig. 5(a), we simulated the autocorrelation traces of the pulses. The simulated interferometric autocorrelation and intensity autocorrelation are shown in Fig. 6(a) and 6(b), respectively. The intensity autocorrelation in Fig. 6(b) shows the width of autocorrelation (FWHM) is 26.6 fs. With the reconstructed pulse waveform in Fig. 5(b) and the simulated intensity autocorrelation in Fig. 6(b), we can obtain the ratio of autocorrelation width  $\tau_c$  to the pulse width  $\tau_p$  is:  $\tau_c/\tau_p = 1.462$ .

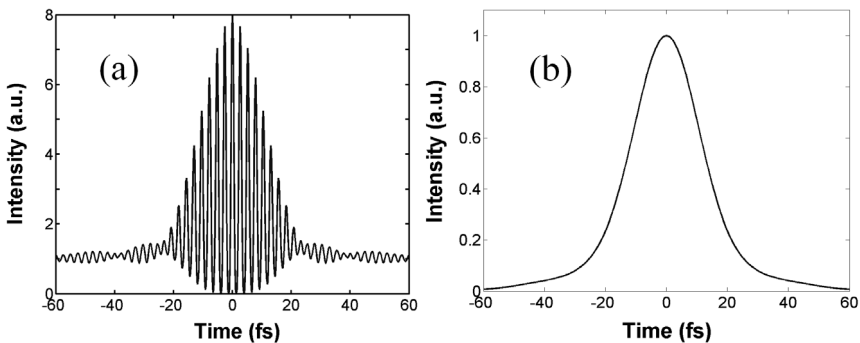
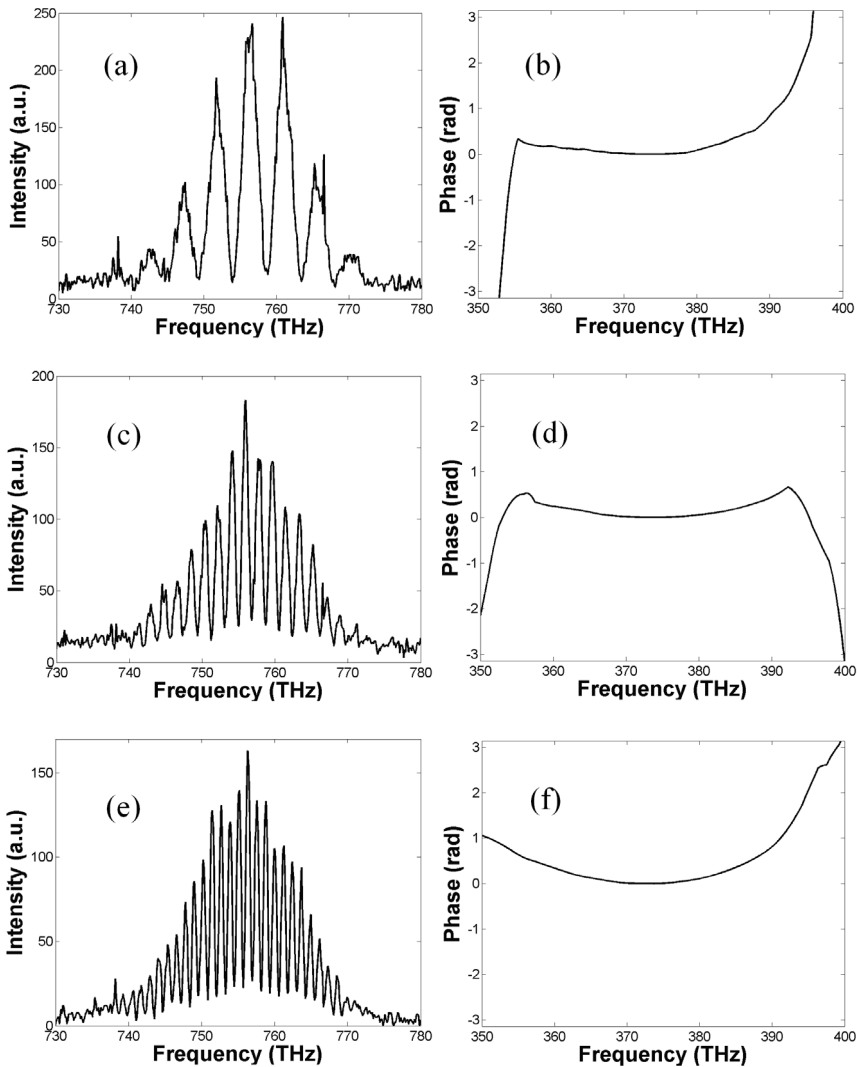


Fig. 6. Simulated autocorrelation traces with reconstructed pulse electric field. (a) Interferometric autocorrelation, (b) intensity autocorrelation.

## 2.4 Pulse waveform reconstruction from different replicas separations

We have investigated the effects of pulse replicas separation on the spectral phase retrieval of femtosecond optical pulses. By tuning the translation stage TS2, a series of spectral interferograms with different replicas separations was recorded. Figures 7(a), 7(c), 7(e), 7(g), and 7(i) are five measured spectral interferograms with replicas separations of 0.22 ps, 0.55 ps, 0.83 ps, 1.58 ps, and 1.82 ps respectively. We retrieved the spectral phases from the five measured interferograms with wavelet-transform. The extracted spectral phases are shown in Figs. 7(b), 7(d), 7(f), 7(h), and 7(j).



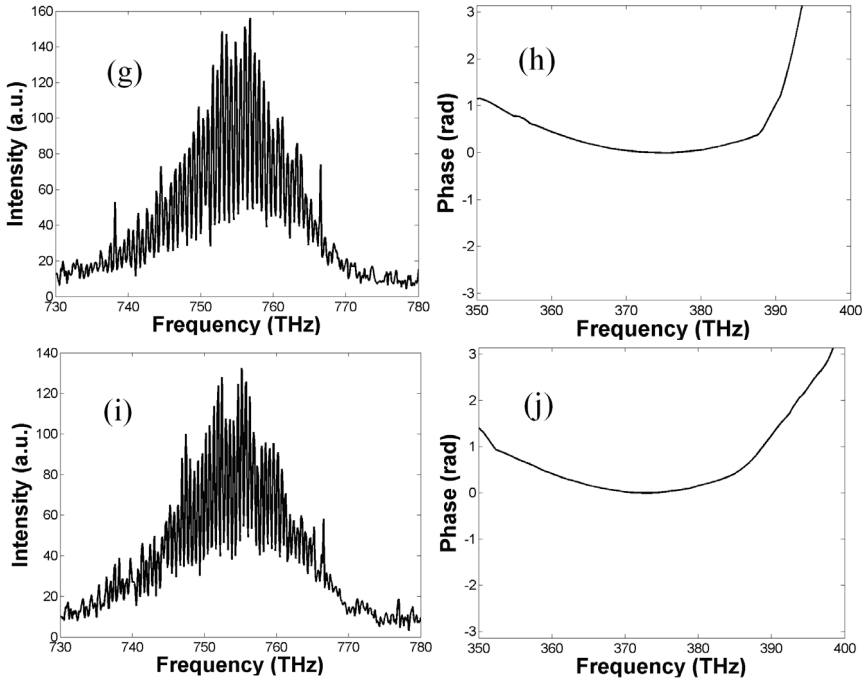


Fig. 7. Measured spectral interferograms with different replicas separations and retrieved spectral phases from the interferograms. (a), (c), (e), (g), and (i) are measured interferograms with replicas separations of 0.22 ps, 0.55 ps, 0.83 ps, 1.58 ps, and 1.82 ps respectively; (b), (d), (f), (h), and (j) are the retrieved spectral phases from interferograms in (a), (c), (e), (g), and (i) respectively.

With the retrieved spectral phases and the measured spectrum, the waveform of the pulse can be reconstructed. Figure 8(a) is the five retrieved spectral phases with different replicas separations and the spectrum. Figure 8(b) is the reconstructed pulse profiles from the spectrum and spectral phases with inverse Fourier transform.

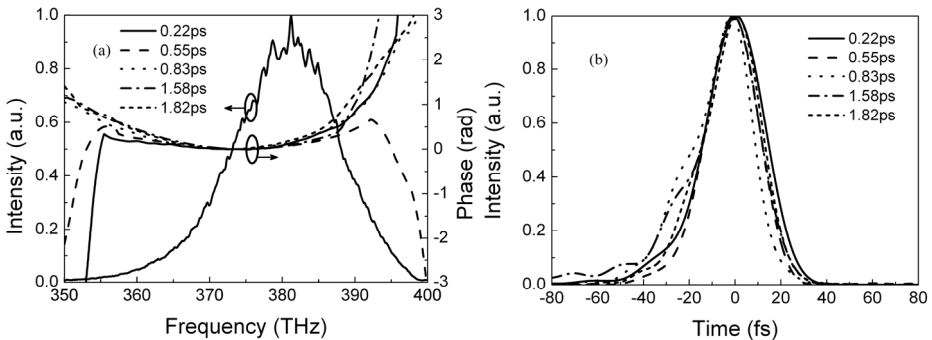


Fig. 8. Retrieved spectral phase and reconstructed pulse profile with different replicas separations. (a) Spectrum and retrieved spectral phases, (b) reconstructed pulse profiles.

Figure 8(a) shows that with wavelet transform, spectral phases are retrieved from a large range of replicas separations. With replicas separations from 0.22 ps to 1.82 ps, the relative difference of spectral phase is within 0.5 rad. In Fig. 8(b), the full width at half maximum (FWHM) of the reconstructed pulse profiles from replicas separations of 0.22 ps, 0.55 ps, 0.83 ps, 1.58 ps, and 1.82 ps are 28.50 fs, 27.30 fs, 27.77 fs, 28.75 fs, 28.47 fs, 26.81 fs, 26.65 fs, 27.60 fs, and 28.03 fs, respectively. The maximum relative difference is 2.76%, which demonstrates the accuracy of the spectral phase retrieval with wavelet-transform.

## 2.5 Uncertainty discussions

In the theory of inverse Fourier transform, reconstruction of pulse profile needs two parameters: spectrum and spectral phase. Broad spectrum of femtosecond optical pulse is measured with a spectrometer, which can be accurately calibrated by a black body or a standard lamp. Therefore, the uncertainty of spectral phase contribute majority of the uncertainty of pulse reconstruction. Uncertainty budget of spectral phase measurement is complex. We assume the laser source runs in a stable status, and the fluctuation of a mode-locked pulse is ignorable. The effects of the temperature and humidity of the surroundings on the chromatic dispersion of the measurement instrument is also ignorable. Therein we simply discuss the three main components for the uncertainty generation: measurement instrument (SPIDER setup), measured spectral interferogram, and algorithm of phase retrieval.

The chromatic dispersion of the beam splitter and the nonlinear crystal of the SPIDER setup generates an additional phase of the measured pulse. This additional phase shows an unnegligible effects in measurement of short pulses, especially in the case of pulse width less than 5 fs. Therefore, the thickness of beam splitters and the nonlinear crystals must be very small. To reduce the spectral phase retrieval error, the optical paths need carefully alignment for a high signal to noise ratio (SNR) interferogram recording. The spectral phase retrieval from interferograms with different replicas separations have been analyzed in section 2.4. The analysis and comparison of spectral phase retrieval error with wavelet-transform and Fourier transform with different filter widths is simulated and calculated in reference [10].

## 2.6 Comparison of wavelet-transform with Fourier transform

The comparison of spectral phase retrieval procedure with wavelet-transform and that with Fourier transform is plotted in Fig. 9. Figure 9(a) is the flow chart of phase retrieval by Fourier transform, and Fig. 9(b) is that by wavelet-transform.

From Fig. 9, we can see that phase retrieval procedure with wavelet-transform is much simpler. The phase reconstruction by Fourier transform includes three steps: Fourier transform, filtering, and inversed Fourier transform. In contrast, phase information can be read directly after the wavelet-transform, which saves two steps in the phase reconstruction, thus simplified the procedure. This simplified process eliminated the uncertainty coming from the filter with traditional Fourier transform; therefore more accurate spectral phase is retrieved. Another advantage of this simplified procedure is that the manual process of selection and adjustment filter is eliminated; therefore, the spectral phase compensation and control can realize automatic operation.

In principle, Fourier transform mixes all the time information of the spectral interferogram at different frequency position no mater whether there are signal or not, therefore, it is more prone to bring noise into signal information and result in pseudo phase.

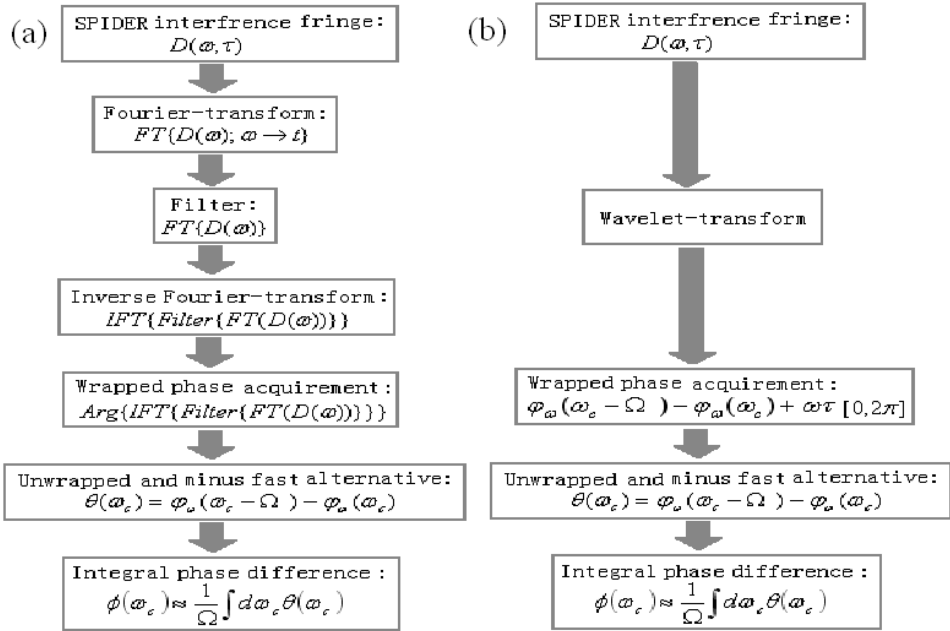


Fig. 9. Comparison of spectral phase retrieval flows between by Fourier transform and by wavelet-transform. (a) Flow chart of phase retrieval by Fourier transform, (b) by wavelet-transform.

However, the major advantage of phase reconstruction using wavelet-transform technique is that the wavelet-transform transfers the spectral interferogram into a two-dimensional graph that separates the signal from noise naturally. It does not need a filter to obtain phase information, but just probe the phase at the maximum of the transform. Therefore the error coming from the filter and the envelope noise is avoided; because there is no filter that needs to be judged by human, all the calculation can be processed automatically and is more suitable for video speed display.

### 3. Characterization of ultrashort optical pulses with autocorrelations

There are mainly two kinds of autocorrelations, i.e., intensity autocorrelations and interferometric autocorrelations [6]. Both of them have a Michelson-type interferometer setup. Intensity autocorrelation uses a translation stage scanning the split pulses to give an intensity envelope of the autocorrelation. The pulse width can be deduced from the width of the autocorrelation envelope. Intensity autocorrelation has a broad measurement range by use of the translation stage; however, it cannot provide complete information about the pulse shape. Interferometric autocorrelation uses a vibratile arm real-time scanning the split



pulses to generate a distinctive autocorrelation trace. The chirp and phase of the ultrashort optical pulse are reflected in the interferometric trace; however, the vibration range is limited in several hundred micrometers, which results in a limitation of measurement range with this technique. Therefore, a multifunctional autocorrelator, which can realize both interferometric autocorrelation and intensity autocorrelation, is desirable.

Meshulach *et al* have demonstrated a third-harmonic generation (THG) autocorrelator for ultrashort optical pulses measurement [12]. They used an ordinary glass slide as the THG source, and a pair of 32:1 THG interferometric autocorrelation trace and intensity autocorrelation trace were obtained. However, THG needs the input pulse have enough high energy, and the generated signal always has a low SNR in weak energy conditions. In this letter, we demonstrate a second-harmonic generation (SHG) autocorrelator. We use a two-photon detector as the source of SHG. The setup is simplified, and the sensitivity is improved.

### 3.1 Optics schematics of a multifunctional autocorrelator

A multifunctional autocorrelator can be realized by use of a precise translation stage and a lock-in amplifier. The optical schematic is shown in Fig. 10(a). BS1 and BS2 are two 30/70 beam splitters coated with 450 nm - 1150 nm broad bandwidth film, and M1 and M2 are two silver roof mirrors. M2 was mounted on a translation stage (M405-DG, Physik Instrument GmbH), the resolution of which can reach 8.5 nm to provide a fine enough optical delay. We place a chopper (SR540, Stanford Research Systems Inc.) in the inlet of the incident pulse, and the output of a two-photon detector (G1115, Hamamatsu Photonics Corp.) is fed into a lock-in amplifier (SR830, Stanford Research Systems Inc.) to improve the SNR. Our home-made experimental setup of the multifunctional autocorrelator is shown in Fig. 10(b).

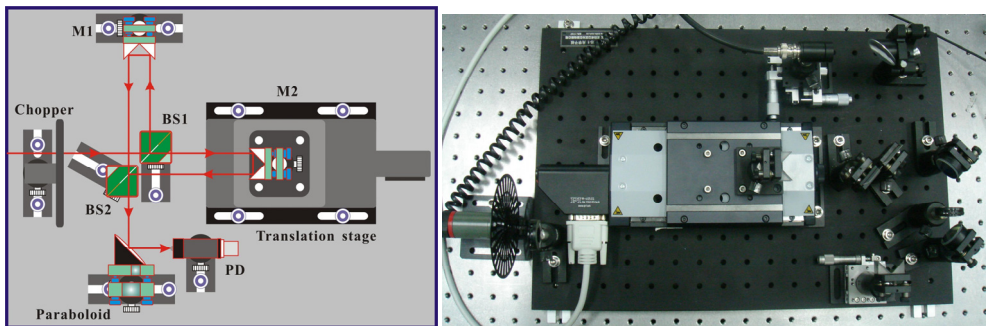


Fig. 10. Schematic of the versatile autocorrelator. M1, M2: Roof mirror; BS1, BS2: Beam splitter. (a) Optical schematic, (b) experimental setup.

The interferometric autocorrelation and intensity autocorrelation can be easily switched by tuning the scanning speed of the translation stage and the time constant of the lock-in amplifier. When the ratio of the central wavelength of the measured pulse  $\lambda_c$  to the time constant of the lock-in amplifier  $t_c$  is much less than two times of the scanning speed of the translation stage  $v_s$ , i.e.,  $\lambda_c / t_c \ll 2 \cdot v_s$ , the measurement result shows an intensity

autocorrelation trace; When the ratio of the central wavelength  $\lambda_c$  to the time constant of the lock-in amplifier  $t_c$  is greatly larger than two times of the scanning speed of the translation stage  $v_s$ , i.e.,  $\lambda_c / t_c \gg 2 \cdot v_s$ , the measurement performs as an interferometric autocorrelation trace.

In section 2.2, we have measured the spectral interferogram of an ultrashort optical pulse train emitted from a Ti:sapphire laser (Micra-5, Coherent Inc.). In this section, we measure the autocorrelations of the same laser pulse with our home-made autocorrelator. By tuning the scanning speed of the translation stage and the time constant of the lock-in amplifier, both interferometric autocorrelation and intensity autocorrelation have been obtained. Figure 11(a) shows the intensity autocorrelation, and Fig. 11(b) shows the interferometric autocorrelation.

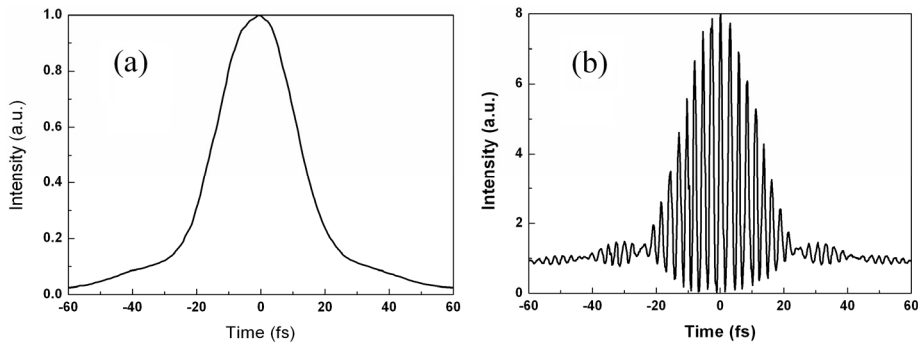


Fig. 11. Measured autocorrelation of Ti:sapphire laser pulses. (a) Interferometric autocorrelation, (b) intensity autocorrelation.

The intensity autocorrelation in Fig. 11(a) shows an autocorrelation width (FWHM) of 30 fs, and the corresponding pulses width is nearly 20 fs. And the interferometric autocorrelation in Fig. 11(b) shows the autocorrelation width and the chirp information. The interferometric autocorrelation has such a high SNR that the two wings of the interferometric trace can be clearly distinguished, which is useful in complete characterization of the ultrashort optical pulses.

From Fig. 11, we can see that the simulated autocorrelations are in good agreement with the simulated ones (in Fig. 6(b) and 6(a)). The deviation is within  $\pm 1$  fs, which gives a proof of the reliability for both the techniques.

### 3.2 Measurement of a strongly chirped pulse

This autocorrelation setup is suitable for characterization of various ultrashort optical pulses, including strongly chirped pulses and fiber laser pulses. We use a 25-mm-thick SF 10 glass bulk to stretch the ultrashort optical pulses emitted from the Ti:sapphire laser. The chromatic dispersion of SF 10 glass is  $156 \text{ fs}^2$  calculated with Sellmeier coefficient. And the width of critical pulse can be gotten with the equation:

$$T_c = 2\sqrt{\ln 2 \times |\phi''|} = 104 \text{ fs.}$$

For a 15 fs nearly transform-limited ultrashort optical pulse, the output pulse width from the SF 10 glass is:

$$t_{p,out} = [1 + (T_c/t_p)^4]^{1/2} \cdot t_p = 720 \text{ fs.}$$

The pulses were stretched in temporal range with a great quadratic chirp, which can also be well characterized with the autocorrelator. The intensity autocorrelation and the interferometric autocorrelation were shown in Fig. 12(a) and 12(b), respectively.

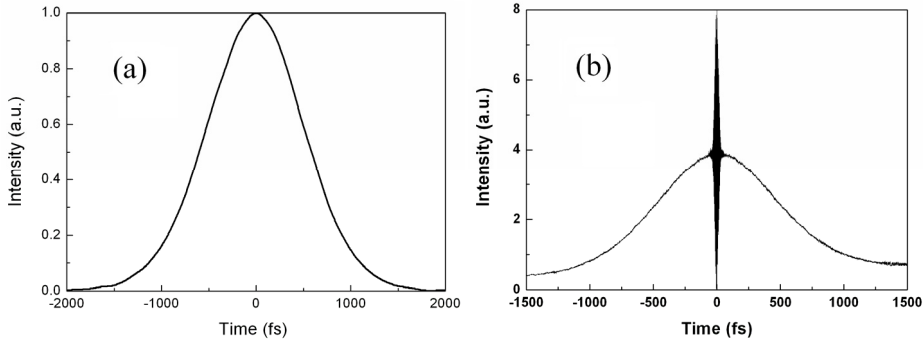


Fig. 12. Measured autocorrelation of the Ti:sapphire laser pulses stretched by a SF 10 glass bulk. (a) Interferometric autocorrelation, (b) intensity autocorrelation.

In Fig. 12(a), the intensity autocorrelation clearly shows the width of the broadened pulses. We can clearly discern the pulses width and the chirp of the pulses with the two autocorrelation traces, even if they have a great chirp. The intensity autocorrelation in Fig. 12(a) shows an autocorrelation width (FWHM) of 1100 fs, and the corresponding pulses width is 720 fs assuming a Sech<sup>2</sup>-shaped pulses, which can agree well with the calculation of the stretched pulse. From Fig. 12(b), we can see that the interferometric autocorrelation shows the interference range is very narrow, and the two wings uplift, which demonstrated the great quadratic chirp generated from the dispersion of the SF 10 glass.

With the intensity autocorrelation and interferometric autocorrelation, both pulse width and phase information of the pulses were presented. Another advantage of this technique is that the space of the interferometric fringes can be used as a ruler to calibrate the time axis of the intensity autocorrelation. That is to say, this autocorrelator setup can self-calibrate with the two autocorrelation traces. This autocorrelation setup can give an accurate autocorrelation width of the pulses; however, a ratio factor should be assumed from autocorrelation width to pulse width. The ratio factor of different pulse waveforms varies greatly; therefore, unsuitable factor may generate great error in pulse width determination with autocorrelation.

### 3.3 Measurement of a fiber laser pulse

The autocorrelator also performs excellently in characterization of fiber lasers pulses. We measured a photonic crystal fiber laser (Ultrafast Laser Laboratory, Tianjin University) [13] with this autocorrelator. The output average power of the laser is 500 mW, central

wavelength is 1060 nm, and pulse repeat frequency is 25 MHz. The intensity autocorrelation and the interferometric autocorrelation traces were shown in Fig. 13(a) and 13(b), respectively.

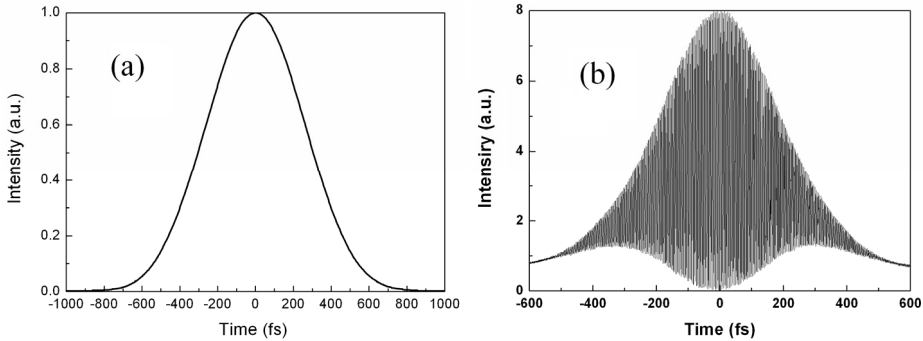


Fig. 13. Measured autocorrelations of photonic crystal fiber laser pulses. (a) Interferometric autocorrelation, (b) intensity autocorrelation.

Figure 13(a) gives a desirable intensity autocorrelation, which can clearly show the pulse width. The intensity autocorrelation applied a fast scanning of the translation stage; it is therefore much timesaving and suitable for fast measurement. The interferometric autocorrelation was shown in Fig. 13(b). The interferometric trace can be clearly discerned even the scanning length has extended to 1200 fs. From Fig. 13(b), we can see the dispersion compensation is acceptable except for a slight quadratic chirp. The interferometric autocorrelation trace in Fig. 13(b) gives phase information of the optical pulses, which is useful for the developing and optimizing of the laser.

This autocorrelator is based on a translation stage; therefore, the measurement range is related to the translation range of the stage. The stage we chosen has a translation range of 50 mm; therefore, the measurement range of the autocorrelator is 110 ps. By use of a two-photon detector and a lock-in amplifier, this autocorrelation has a higher sensitivity than THG one. The pulses with a 10 mW average power can be well measured.

#### 4. Conclusions

The characterization technique of ultrashort optical pulses is introduced and demonstrated in this chapter. For accurate characterization of femtosecond optical pulses, we introduced a wavelet-transform technique for spectral phase retrieval. The wavelet-transform technique needs no filter; therefore, automatic and accurate phase retrieval and pulse reconstruction is realized because of no manual operation of selection and adjustment filter. With this technique, the ratio of autocorrelation width to pulse width can be determined, which is helpful in improving the precision of pulses width measurement with autocorrelation technique. We have demonstrated a multifunctional autocorrelator, based on a precise translation stage and a lock-in amplifier, for pulse width measurement of ultrashort optical pulses. With the autocorrelator, both interferometric autocorrelation and intensity autocorrelation can be obtained. Chirp can be clearly discerned from interferometric autocorrelation, and pulse width can be fast obtained from intensity autocorrelation. The

measurement result can self-calibrate with the two correlation traces. The autocorrelator is suitable for the measurement of ultrashort optical pulses from sub-10 fs to several hundred picoseconds, and it is useful in generation and characterization of both solid ultrashort pulse lasers and fiber ultrashort pulse lasers.

## 5. Acknowledgment

This work was supported in part by the National Science and Technology Supporting Program of China (grant 2006BAF06B05), Basic Research Foundation of National Institute of Metrology, China (grant AKY0748, AKY0904, and AKY1160).

## 6. References

- [1] Ell R., Morgner U., Kärtner F. X., Fujimoto J. G., Ippen E. P., Scheuer V., Angelow G., Tschudi T., Lederer M. J., Boiko A., Luther-Davies B. (2001). Generation of 5-fs pulses and octave-spanning spectra directly from a Ti:sapphire laser, *Opt. Lett.*, Vol. 26, pp. 373-375
- [2] Margolis H. S., Harper M. R., Lea S. N., (2006). Metrological applications of femtosecond optical pulse shaping, *NPL Report*, DEM-EM-010
- [3] Hall J. L. and Ye J. (2001). A new era of frequency standards and optical frequency measurement, *Optics & Photonics News*, Vol. 12, pp. 44-50
- [4] Yasui T., Yokoyama S., Inaba H., Minoshima K., Nagatsuma T., and Araki T. (2011). Terahertz frequency metrology based on frequency comb, *IEEE J. Sel. Top. Quant.*, Vol.17, pp. 191-201
- [5] Williams D. F., Hale P. D., Clement T. S., and Morgan J. M. (2005). Calibrated 200-GHz waveform measurement, *IEEE T. Microw. Theory.*, Vol.53, pp. 1384-1389
- [6] Diels J.-C. M., Fontaine J. J., McMichael I. C., and Simoni F. (1985). Control and measurement of ultrashort pulse shapes (in amplitude and phase) with femtosecond accuracy, *Appl. Opt.*, Vol.24, pp. 1270-1282
- [7] Kane D. J., Trebino R. (1993). Single-shot measurement of the intensity and phase of an arbitrary ultrashort pulse by using frequency-resolved optical gating, *Opt. Lett.*, Vol.18, pp. 823-825
- [8] Iaconis C. and Walmsley I. A. (1999). Self-referencing spectral interferometry for measuring ultrashort optical pulses, *IEEE J. Quantum Electron.*, Vol.35, pp. 501-509
- [9] Kärtner F. X. (2005) Chapter 10: Pulse Characterization Ultrafast Optics (Spring Term 2005), pp. 333-370
- [10] Deng Y., Wu Z., Cao S., Chai L., Wang C. and Zhang Z. (2006). Spectral phase extraction from spectral interferogram for structured spectrum of femtosecond optical pulses, *Opt. Commun.* Vol.268 pp. 1-6
- [11] Deng Y., Wu Z., Chai L., Wang C., Yamane K., Morita R., Yamashita M. and Zhang Z. (2005). Wavelet-transform analysis of spectral shearing interferometry for phase reconstruction of femtosecond optical pulses, *Opt. Express* Vol.13, pp. 2120-2126
- [12] Meshulach D., Barad Y., and Silberberg Y. (1997). Measurement of ultrashort optical pulses by third-harmonic generation, *J. Opt. Soc. Am. B*, Vol.14, pp. 2122-2125
- [13] Song Y., Hu M., Wang C., Tian Z., Xing Q., Chai L. and Wang C. (2008). Environmentally Stable, High Pulse Energy Yb-Doped Large-Mode-Area Photonic Crystal Fiber Laser Operating in the Soliton-Like Regime, *IEEE Photon. Tech. Lett.*, Vol. 20, pp. 1088-1090

## **Section 5**

### **Time and Relativity**



# Time Measurement and Earth Rotation

Victor Tissen, Alexander Tolstikov and Galina Simonova  
*SNIIM, SSGA*  
*Russia*

## 1. Introduction

This chapter covers some issues of measuring, storage and reproduction of a precise time scale and reference frequencies, as well as techniques and methods for providing consumers with this scale and Earth rotation parameters (ERP) data.

The first paragraph contains historical information on development of time measurement and storage instruments. It reviews development stages of national and world time services, their duties and goals within the scope of an assurance of time and frequencies measurement and reproduction uniformity.

The second paragraph represents a brief classification of time measurement systems basing on different natural periodic process, which are used as a standard time unit. Possibilities of use a pulsar time scale for long-time storage of high-stable time intervals and synchronization of group time and frequencies references are also reviewed in this paragraph.

The third paragraph states the main objectives of national time and frequencies services, describes a drawing scheme for a group time scale, its generation algorithms and a structure of a measurement instruments complex for time and frequency storage and reproduction.

The fourth paragraph studies methods for synchronization of reference clocks, which are distant from each other making a direct connection between them impossible. Different types of radio and wire communications systems used as solutions for providing various consumers with standard time signal are also described briefly. For example, there are following such consumers: communications and telecommunications, information technologies computer systems and computer servers, space and ground-based navigation, space geodesy, geodynamics, transport management systems, etc. Some synchronization methods based on the use of GPS/GLONASS satellite groups are presented.

The fifth paragraph covers algorithmic and software tools, which allow to evaluate the accuracy of solving of time and coordinates support problems by means of a simulation modelling. In particular, a brief description of the simulation model of the quantum clock instability developed by the authors and used in a software simulator of the measurement data, received from a network of non-query measuring stations through the GLONASS and GPS navigation satellites is given in this paragraph.



A comparative analysis based on the main metrological characteristic of the modern methods of ERP determination and prediction is conducted in the sixth paragraph. The ERP prediction accuracy requirements for the time intervals of the coordinate and ephemeris support of GNSS operation are given within the scope of metrological support issues of coordinate and navigational determinations for GPS and GLONASS GNSS. The basic principles and approaches used by authors for the development of new high-precision ERP prediction method are set forth.

## 2. History of information

Need in time measurement arose about 5000 years ago, around a transition from nomadic to sedentary lifestyle, farming and cattle breeding. In particular, need in observance of optimum time for agricultural work and coordination of collective actions caused invention of different methods and devices for long and short time period measurement that were based either on uniform or periodic natural processes. Diurnal solar motion started to be used for time measurement in Ancient Egypt about 3500 B. C. Obelisks of a certain height ("gnomons") were installed in strategic locations to mark noon with the shortest shadow. Two thousand years later the Egyptians invented sun-dial, "face" of which was divided into 10 parts. By the beginning of the Common Era in Mediterranean industrialized countries more than 30 kinds of sun-dials were used. (Mihal, 1983). In order to measure short time periods, such uniform natural processes were used as flow of liquid or light sand through a narrow hole. Sand and water clocks based on this principle were widespread in the ancient world and sometimes are still used nowadays.

First stellar work for the purpose of time measurement began about 600 years B. C. by means of the astronomical tool "merkhet" (<http://www.infoniac.ru>). After mechanical spring-actuated and pendulum clocks were invented in the early XIV century A. D. in Europe, precision of time period measurements considerably improved. Pendulum clock constructional design was improved by Galileo, Huygens and Hooke. On the basis of their work in the XVIII century wrist watch was made, it was 5 seconds in error and its manufacture started in the XIX century. (Pipunyro, 1982).

The XX century was characterized by huge scientific and technical progress in terms of mastering new time measuring methods and devices. First in 1920 by Short, and then in 1955 by Fedchenko the best astronomical pendulum clock constructional designs were developed, they were 2ms/per day and 0.2 ms/per day in error respectively. (Bakulin & Blinov, 1977). Quartz-crystal clock, invented by Morrison in 1927 and based on the piezoelectric effect, opened up new possibilities for improvement of time measurement and time keeping. By the mid-1950s time scale, determined by thoroughly investigated quartz crystal clock, started to serve as an independent time standard, more stable than diurnal rotation of the Earth. As a result, the possibility to detect instability of earth time scale first appeared, as well as the possibility to detect nonuniformity of the Earth's rotation. But quartz crystal clock grave disadvantage was that they could not maintain stability over long time periods because of quartz ageing effect (Great Russian Encyclopedic Dictionary [GRED], 2003).

A complete swing-round in terms of time keeping occurred after atomic and molecular frequency standards were invented. In 1955 the Englishman Essen developed the first atomic frequency (time) standard by means of a cesium beam. From that moment on atomic

second, equivalent to a time span over which a cooled caesium atom emits 9192631770 radiation periods, is considered to be the world time and frequency standard. Atomic time standard is free from both diurnal and secular variations, does not have ageing effect and is characterized by sufficient determinateness, precision and reproducibility. With the use of atomic frequency standards, time standards independent from the Earth's rotation were developed, they are characterized by exceptionally high stability over long time periods. It enabled to fulfil the task of large distance measurement metrological assurance. Precise time measurements are also required to fulfil different tasks of navigation, telecommunications, terrestrial and extraterrestrial navigation, geodesics and geodynamics. They are also used in traffic control systems, information technologies and other spheres. A series of basic researches – aimed at a more precise definition of principal laws of nature related to extension of knowledge concerning Macrocosm, Aerospace, Earth, Microcosm – require such a precise standard that approaches to a limit determined by fundamental physics laws. Thus far the highest level of relative precision of playback of atomic second of the order of  $3 \cdot 10^{-16}$  was reached by the American National Institute of Standards and Technology (NIST) by means of caesium fountain clock. In the years to come, a new generation of frequency standards that radiate frequencies not in microwave but in optical spectrum is expected. Optical clock is developed especially actively in Japan. For example, Hidetoshi Katori's group of Tokyo University has already reached precision of the order of  $10^{-15}$  by means of an experimental model of strontium optical clock, and theoretically they can maintain precision of playback of time and frequency units at  $10^{-17} \div 10^{-18}$  (Hall, 2006).

### 3. Time measurement systems

Time measurement is one of the most important tasks of modern metrology, astrometry and physics. In order to fulfil this task, it is necessary:

1. to determine time measurement units and time keeping systems;
2. to have at one's disposal reasonably designed time-interval recorders;
3. to control operation of these recorders either occasionally or constantly.

It is easy to use periodic processes based on natural life cycles and characterized by high stability over long time periods as standard time units. Among all processes used for determining time units there are three basic ones:

1. the Earth's axial rotation;
2. the Earth's circuit of the Sun;
3. emission (absorption) of electromagnetic waves by atoms and molecules of certain substances under certain conditions.

Thanks to these natural processes, three independent time scales were developed; these are universal (astronomical), ephemeris and atomic time scales. Universal or worldwide time scale was so called as it is directly connected to the nychtemeral cycle that is the main cycle of people's lives. In terms of a type of a process that forms a universal time scale, there are three universal time systems.

1. *UT0* is universal time registered directly as a result of astronomical observations. Technically, it is not universal, as it depends on where on the Earth the observatory is located.

2. *UT1* is universal time which was corrected as a result of change of longitudes of observatories under the influence of movement of poles, it is calculated according to the formula:

$$UT1 = UT + \Delta\lambda, \quad (1)$$

where  $\Delta\lambda$  is a correction to the observatory longitude, for movement of the North Pole.

3. *UT2* is universal time, which also takes into account seasonal variations in the Earth's rotation speed. *UT2* is the most uniform time, which can be received from astronomical observations of diurnal rotation of the Earth. *UT2* and *UT1* time scales differ in the value of seasonal variations  $\Delta T_s$ :

$$UT2 = UT1 + \Delta T_s. \quad (2)$$

The ephemeris time scale was introduced in 1952 by the decision of Paris international conference on fundamental astronomical constants. Ephemeris time (*ET*) is an independent variable of celestial mechanics equations and is defined as difference between observed and calculated, according to celestial-mechanical theories, coordinates of the Moon, the Sun and planets. As *ET* standard a second, equal to  $1/31556925,9747$  part of a tropical year, was taken. Uniformity of *ET* scale was influenced by observation errors and uncertainties of adopted theories of the Moon's and planetary movement (Bakulin & Blinov, 1977).

In 1967 by the decision of the XII General Conference of Weights and Measures the atomic time scale *TAI* was introduced, the scale unit is equal to a time span, in which the caesium atom goes through 9192631770 emission periods. The atomic scale is characterized by the highest stability of all mentioned ones and that is why it is the main time standard in the world. It is connected to the ephemeris time scale by the formula:

$$ET = TAI + 32,18 c \quad (3)$$

Universal Coordinated Time scale *UTC* is a combined one; it is based on the atomic time scale, indications of which are corrected taking into account data concerning the Earth's rotation. From time to time *UTC* indications are corrected to 1 second, so that difference between *UT1* – *UTC* does not exceed 0,9 second in modulus. *UTC* signals are transmitted in broadcasting networks and used in daily life. The *UTC* advantages are high uniformity that is characteristic for atomic time, and connection to natural processes (sunrise, sunset) that is characteristic for solar time. Some countries form and maintain their own *UTC* scale. For example, in the U.S. the *UTC* (USNO) scale is based on the assembly of (about 50) caesium standards, its indications are not more than 50 ns in error compared to the international standard *UTC*. In Russia the national scale *UTC* (SU) is maintained which is offset by +3 hours from *UTC* (Belotserkovsky & Kaufman, 1972).

The possibility of independent pulsar time scale formation has lately been the subject of wide speculation. Pulsars are speed-up neutron stars that represent sources of high-stable radio pulses with frequencies of the order of 1-1000 Hz. In view of this property, pulsars are regarded as potential time keepers, located in extraterrestrial environment. This remarkable property of pulsars is especially important in view of creation and development around the globe of Global Positioning Systems that are based on metric unity of four-dimensional

space-time continuum. (GRED, 2003). But there are natural factors that influence radio-frequency radiation stability of pulsars and are related to their inherent properties and difference in conditions of signal transmission in interstellar medium. So, for actual use of pulsars as time keepers it is necessary to minimize accidental variations and systematic deviations of observed radio frequencies. This is achieved by formation of a pulsar time group scale and application of parameterization methods. In the long term integration of pulsar time with a group atomic keeper for the purpose of atomic time scale correction will enable to carry out comparative measurement and determine corrections to bring indications of all atomic keepers in line with high-stable pulsar time scale, using it as the key standard. In European Radio Astronomical Observatories systematic observations of 25 pulsars are carried out. In the U. S. observation of an assembly of 25 pulsars are carried out as part of the NANOGrav program (North American Nanohertz Observatory for Gravitational Waves) by means of the largest radio telescopes: Arecibo (300m) - 13 MP, RT-100 etc. (Ilyasov and others, 2010).

#### **4. Facilities and methods of time unit and frequency playback in time services**

##### **4.1 Basic tasks of time services**

Basic tasks of exact time services are determination, storage and transmission of standard time signals, as well as determination and forecasting of the Earth's rotation parameters (ERP). There are primary and secondary time standards and frequencies, as well as reference standards, in some countries. A national primary time and frequency standard (NTFS) is designed for ensuring unity of measurements to a required degree of precision, carried out by time services of the country. This is achieved by a number of activities, and the main ones are as follows:

- playback and storage of national time scales;
- coordination of activities that are carried out in centres for metrology services and are aimed at ensuring operation of secondary time and frequency standards;
- comparison of NTFS clocks' indications with standards of other countries;
- control of standard time and frequency signals' (STFS) transmission by means of different technical devices;
- collection and processing of time-and-frequency information provided by domestic time services;
- delivery of necessary reference information to interested organizations and STFS consumers by way of real-time data dissemination and of official bulletins' issue with results of NTFS application;
- Storage of time and frequency units' sizes and maintaining of the coordinated time scale *UTC*, that is brought to the maximum in line with units and a coordinated time scale of national standards.

Dissemination of standard time and frequency is carried out by transmission of time and frequency signals over broadcast channels, by means of GNSS, as well as on the Internet. One of the main tasks of a time service is ensuring the possibility of signal reception anytime anywhere in the required area. Time precision in case of radio transmission is about 1 ms, transmission over television channels - 10  $\mu$ s, over global positioning systems GPS/GLONASS - 10 ns, on the Internet - from 10 ms to 10 ns.

## 4.2 Scheme of group time scale formation and storage

Scheme of exact time units' formation with transmission to consumers over broadcast channels and the Internet is given in Figure 1 by the example of Novosibirsk time service.

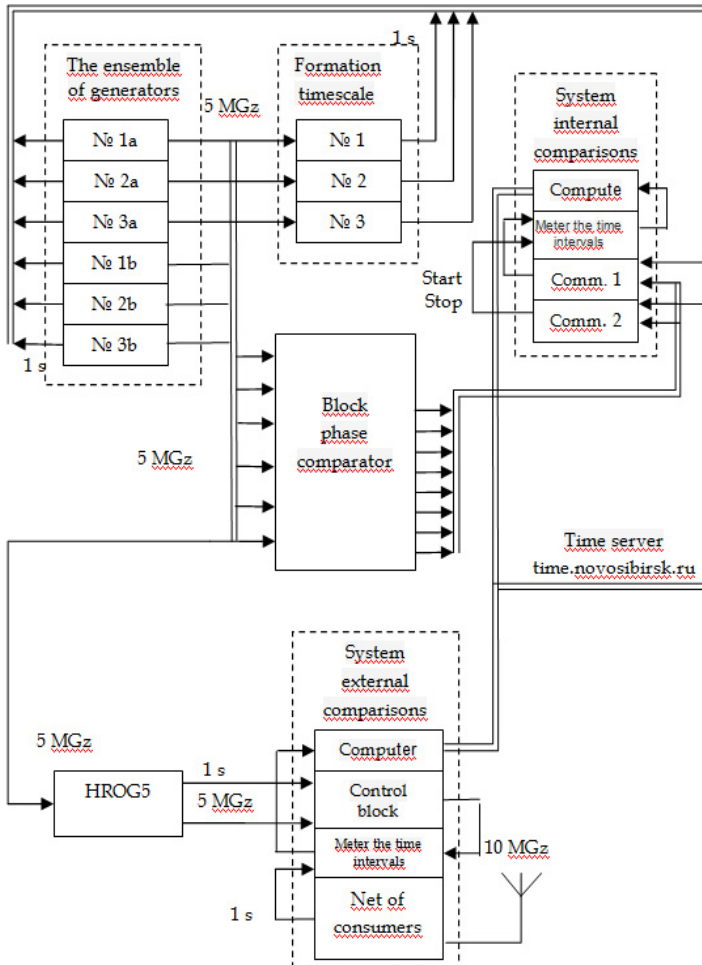


Fig. 1. Block scheme of exact time scale formation equipment

A Standard's atomic time scale is realized by a group keeper which includes at least four hydrogen standards, one of which is selected as a key one. In order to calculate the scale, results of mutual comparison of the Standard keepers' frequency and time scales are used, this is internal comparison. The key keeper is compared to the national standard over communication channels. The results serve as the basis for forecast of frequency variations of the Standard's key keeper. Hereafter, the forecast data are used only in view of routine intercomparison of keepers' time scales. In practical work maintaining of the Standard's atomic scale comes down to determination of time corrections for a key hydrogen keeper.

Frequency and time group keeper is designed for generation of sinusoidal signal with standard frequency of 5 MHz.

Group keeper equipment includes:

- Six hydrogen frequency and time standards of Ч1-75A type;
- Time scale formation system (TSFS), ensuring formation of a group high-precision, high-stable sinusoidal signal with frequency of 5 MHz ;
- a block of phase comparators, which compare signals from each pair of hydrogen standards for time scale correction according to internal comparisons;
- signal switching systems that form control signals for a meter of time intervals;
- time intervals' meters ChZ-64, SR-620, by means of which internal comparisons of differences between scales and individual hydrogen standards are carried out;
- internal comparisons' electronic system that is used to control equipment and transmission of the formed time scale to net of consumers.

Basic metrology characteristics:

- Range of standard frequencies: 1 Hz, 5 MHz, 10 MHz, 100 MHz, 2.048 MHz;
- fractional error of reproduction of a unit of frequency in the form of root-mean-square error measurements' results under condition of daily comparisons with the Russian National time and frequency standard does not exceed  $5 \cdot 10^{-13}$ ;
- difference between time scale *UTC* (NOV) and time scale *UTC* (SU) of Russia does not exceed  $\pm 100$  nc;
- Mean square fractional two-sample deviation (Alan variations) of the group keeper frequency in intervals size of which is from 1s to 1 day are given in Table 1.

Item No.	$\tau_m$	$\varepsilon_m$
1	1 sec.	$3 \cdot 10^{-13}$
2	10 sec.	$5 \cdot 10^{-14}$
3	100 sec.	$1 \cdot 10^{-14}$
4	1 hour	$5 \cdot 10^{-14}$
5	1 day	$1 \cdot 10^{-14}$

Table 1. Alan variations of the group time keeper frequency

Transmission of exact time signals to consumers is carried out on the INTERNET by means of NTP protocol which ensures that error related to *UTC* (NOV) scale is not more than several milliseconds. At the same time, time and frequency signals transmitted by radio stations of short-wave, medium-wave and long-wave range and global positioning systems GLONASS and GPS are controlled.

### 4.3 Algorithms of group time and frequency scale formation

Grouping of keepers ensures high precision and reliability of autonomous storage and playback of group time and frequency scales in intervals of long duration. This is achieved by application of the corresponding algorithms of processing of results of internal and external comparisons of group keeper's time scales. Principle of algorithms is in parameter

estimation of individual keepers according to results of scale and frequency measurements within a group and drift compensation from group time scale nominal value. Changes of points of time of keepers' time scales  $T_i(k)$  and frequency scales  $f_i(k)$ , that form a group, are represented by mathematical models of the kind:

$$T_i(k+1) = T_i(k) + \int_{t_k}^{t_{k+1}} f(\tau) \cdot d\tau, T_i(0) = T_{i0}, \quad (4)$$

$$f_i(k+1) = f_i(k) + d_i(k) \cdot [t_{k+1} - t_k] + w(k+1), f_i(0) = f_{i0}. \quad (5)$$

Within a group of keepers, differences between points of scales  $z_{0i}(k) = T_0(k) - T_i(k) + v_{0i}(k)$  and differences between frequencies  $z_{f0i}(k) = f_0(k) - f_i(k) + v_{f0i}(k)$  with reference to the  $T_0(k)$  scale and the  $f_0(k)$  frequency of the key keeper which has the best characteristics in the group are measured. Estimated time scale point  $\hat{T}_0(0)$  and estimated frequency  $\hat{f}_0(0)$  of the key keeper for the point of time  $t_0$ , received as a result of comparison of the key keeper with time and frequency standard, are assumed as initial conditions for group scale formation.

Algorithm of group keeper's time scale formation includes forecast calculation of group keepers frequencies:

$$\bar{f}_i(k+1) = \hat{f}_i(k) + \hat{d}_i(k) \cdot [t_{k+1} - t_k] \quad (6)$$

according to calculated estimations of frequencies  $\hat{f}_i(k) = \hat{f}_0(k) - z_{f0i}(k)$  and calculation of an average estimation of the key keeper frequency for the ensemble of keepers:

$$\hat{f}_0(k+1) = \frac{1}{n-1} \sum_{i=1}^{n-1} \beta_i \cdot [\bar{f}_i(k+1) - z_{f0i}(k+1)]. \quad (7)$$

The received estimations of frequencies  $\hat{f}_i(k)$  and time scale estimations  $\hat{T}_i(k)$ , calculated according to equations  $\hat{T}_i(k) = \hat{T}_0(k) - z_{0i}(k)$ , are used for calculation of time scales forecast:

$$\bar{T}_i(k+1) = \hat{T}_i(k) + \hat{f}_i(k) \cdot [t_{k+1} - t_k], \hat{T}_i(0) = \hat{T}_{i0}. \quad (8)$$

On the basis of these forecasts and results of scale measurements  $z_{0i}(k+1)$  an average estimation of the key keeper's scale for the ensemble of keepers is calculated:

$$\hat{T}_0(k+1) = \frac{1}{n-1} \sum_{i=1}^{n-1} \gamma_i \cdot [\bar{T}_i(k+1) - z_{0i}(k+1)]. \quad (9)$$

The given algorithms are used in practical work for formation of main time and frequency standards scales.

## 5. Synchronization of spaced-apart clocks

### 5.1 General information

Time scale synchronization implies correction of electronic clock that is correction of its indications and of reference generator frequency. After synchronization is carried out, for

some finite time interval a period of an event at the scale being formed coincides with the required precision with a period of the same event at the standard scale. As time goes, coincidence precision decreases as a result of different disturbing influences on the electronic clock; that is why spaced-apart clocks (comparison of clocks) should be synchronized regularly.

Until recently the most widespread methods of exact time scale synchronization were those that used radio signals of broadcasting stations having transmitted information concerning points and intervals of time at different frequencies. In order to transmit this information to a consumer, different codes and train of impulses can be used; the most widely known among them is the method of 6 points. A transmitting station pings 6 second signals (points) and one extended signal (hyphen) at the beginning of a minute within a system of universal coordinated time *UTC*, and according to the latter by means of receiving equipment time services clocks are synchronized. Precision of synchronization over radio signals is limited by conditions of radio-wave propagation in the earth's atmosphere and by instability of signal delay in receiving and transmitting equipment paths. They are used nowadays mostly in ship navigation, where permissible errors of clocks' indications are several dozens and hundreds of milliseconds. Use of television channels, man-made Earth satellites, as well as natural radiation sources, periodic electromagnetic oscillations from outer space (of pulsars) enables to considerably improve precision of transmission of time and frequency units' sizes.

## 5.2 Tools and methods for synchronization of the time and frequency scale standards

One of the main tasks of the coordinated time software is a support and development of a group standard time and frequency which forms a single group time scale with the possibility of prompt access to it. In order to form a unified group scale there can be used various methods of synchronization of standard time and frequency which are remote from each other by the distances of hundreds and thousands of kilometers. Currently the most accurate and commonly used methods are:

- Transportable quantum clock (TQC);
- Double frequency phase receiving equipment of global navigation satellite systems (GNSS);
- Ground and space communications facilities over broadcasting channels;
- Equipment for duplex comparisons over satellite channels TWSTFT.

Equipment complex for duplex comparisons of time scales over TWSTFT channels (Two-way Satellite Time and Frequency Transfer) ensures synchronization of time scales with an error that does not exceed  $\pm 2$  ns for any geographically distributed time and frequency standards. High precision of this method is explained by the fact that asymmetry of bi-directional channel delay is much less than delay factor for signal propagation in one direction. According to the data given by (Ryzhkov and others, 2007), in duplex method there is no direct correlation between an error and a measuring channel length.

The highest degree of precision in terms of comparison of time scales up to 1 ns is achieved by TQC method, as time scales are compared in close vicinity by means of direct wire connection. Precision of synchronization in TQC method is caused mostly by instability characteristics of transportable clocks and influence of relativistic effect, which appears



because the clocks being compared are located in different coordinate frames, that move in parallel with each other with variable speed and which are different in gravity field potential values. That is why it is also necessary to take into account corrections for relativistic effect in any other methods of synchronization of scales of spaced-apart clocks. Synchronization of time scales by means of TQC, by one- or two-way radio transmission methods, as well as methods of taking into account corrections for relativistic effect and other natural factors which influence precision of measurements, are described in details in the book (Oduan & Gino, 2002).

Synchronization of clocks in compliance with GNSS signals is based on receiving by aerials of spaced locations of the same navigation signals from radiating aerials located in outer space. GNSS satellites radiate exact time signals by means of atomic clock that is synchronized with system time of central synchronizer.

Receiver clock is also synchronized with system time, and as a result it is possible to calculate distance from satellite to receiving aerial according to a measured difference between radiation time and time of receiving signals from the satellite. In order to determine three spatial coordinates of a consumer, delay factors for radio signal propagation from at least three satellites at the same time are measured; it enables to receive a single-valued solution of photographic intersection. In order to determine difference between onboard clock scale and ground borne clock scale, it is necessary to measure additionally a delay factor from one more satellite. Then system of four equations that has 4 indeterminants (three coordinates and difference between scales) and a single-valued solution can be set up. In practical work many redundant measurements from ground-station network by GNSS spacecrafts constellations are carried out. Number of redundant equations achieves hundreds and thousands. These equations are solved with an ordinary least square method and its versions. Uncertainties of measurements at any distances can be reduced to several nanoseconds by averaging for more than one day. Precision of synchronization in this case is influenced most of all by uncertainties of coordinate assignment of receiving aerials and equipment delays in the receiver. That is why coordinates of receiving aerials of non-requesting measuring stations (NMS) and coordinates of spacecrafts (SC) should be known with an error that does not exceed several dozens of centimetres in the same coordinate system. For measurement of delays navigation signals' simulators, which form signals similar to those that are radiated from the spacecraft, are used. In practical work task of synchronization is reduced to calibration of difference between delays of receivers by means of transportation of a standard receiver. Precision of synchronization is also greatly influenced by instability of clock scales located on board a spacecraft and NMS clock scales. Consequently, desynchronization of clock scales occurs which can be partially taken into account by means of an appropriate mathematical model, which would enable to calculate values of compensating corrections. Precision of calculation of these corrections is influenced by:

- errors of an unpredictable deviation of standard clock scale in time intervals between synchronization sessions;
- errors of measurements of points of standard clock scale and spaced-apart clock NMS scale;
- errors of an unpredictable deviation of NMS clock scale in time intervals between synchronization sessions.

According to researches described in the work (Unoshev, 1983), the highest precision of synchronization of spaced-apart clocks A and B is achieved by application of quasi-synchronous receiving method in locations A and B of signals from one SC<sub>i</sub> (method in common view (CV)). This method enables to minimize the listed errors thanks to choice of optimum viewing conditions and application of processing algorithm. Principle of method can be explained by means of analysis of equation of signal propagation delay measurement in the way from SC to consumer equipment (CE) in location A. Let us represent the delay equation as follows:

$$\tau_{iA} = \frac{|\bar{r}_{iA}|}{\nu} + \Delta T_A - \Delta T_i + \delta_{U_{iA}} + \delta_{T_{iA}}, \quad (10)$$

where

$$|\bar{r}_{iA}| = \sqrt{(x_A - x_i)^2 + (y_A - y_i)^2 + (z_A - z_i)^2} \quad (11)$$

is a geometric range from SC<sub>i</sub> to CE in location A;

$\nu$  is average radio signal propagation speed from SC to NMS;

$\Delta T_A = T_{RA} - T_C$  is difference between receiving equipment clock scales in location A and system scale  $T_C$  ;

$\Delta T_i = T_i - T_C$  is difference between SC<sub>i</sub> clock scale and system scale  $T_C$  ;

$\delta_{U_{iA}}$  is signal delay in ionosphere on the way SC<sub>i</sub>-A;

$\delta_{T_{iA}}$  is signal delay in troposphere on the way SC<sub>i</sub>-A.

Let us assume that in locations A and B time keepers are located. At these locations the reception of the same signal from some SC<sub>i</sub> is carried out at the time points  $T_A$  and  $T_B$  and differences between receiving equipment scales  $T_{RA}$ ,  $T_{RB}$  and time keepers scales are measured:

$$\Delta_A = T_{RA} - T_A, \quad \Delta_B = T_{RB} - T_B \quad (12)$$

Then, after data exchange  $\Delta_A$  и  $\Delta_B$  in locations A and B one can find the second differences:

$$\Delta_{AB} = \Delta_A - \Delta_B = e_\Sigma - (T_A - T_B). \quad (13)$$

These differences contain information concerning discrepancy between keepers scales  $T_A$  and  $T_B$  and overall synchronization error  $e_\Sigma$ . The lowest level of synchronization error in the "CV" mode is achieved at the moment SC crosses "traverse plane" (TP) - a plane that crosses the midpoint of segment AB transversely to it.

Now the various consumers set the increasingly higher requirements for time standards synchronization and frequency nominal coincidence. For example, a synchronization of the primary time standards with an accuracy of (100-10) ps and a frequency coincidence of  $10^{-15}$  to  $10^{-16}$  per day is required for an effective operation of the high-speed digital fibre optic communication lines, which transfer data at the rates of tens or hundreds tera-bits per

second. Only the VLBI technology can fulfil such requirements at long distances 2000 km or more (Finkelshtein, A. (2007).

## 6. Simulation of navigation measurements

### 6.1 General information

It is worth using software simulators of navigation measurements for development of model systems designed for registry of effects of various natural factors on spacecraft (SC) GNSS traffic and for improvement of navigation measurement information processing technology. Such simulators help to solve the task of parameter definition of a SC GNSS traffic mathematical model, as well as the task of some nuisance parameter part definition. They also serve to determine metrological estimations of effects of these parameters on navigation measurements results. In order to solve these tasks, a software simulator should fulfil the following functions:

1. computation of motion of a navigation satellites orbit group under conditions that produce an effect of perturbation on satellites;
2. building of non-requesting measuring stations (NMS) network and observed SC constellation;
3. estimation of a geometric range from stations to a SC;
4. imitation of factors that influence precision of trajectory measurements.

Fulfilment of the first item is a theoretical task related to application of the equation of SC traffic in the Earth's gravitational field for the purpose of determination of the selected SC coordinates expected values at preset time points. The second and the third items set parameters describing conditions and an observing station. The fourth item solves tasks of nuisance factors simulation on the basis of the navigation measurements equation:

$$S_{ii} = D_{2i}(X_{KA}(t_{2i}), X_{MII}(t_{3i})) + \delta D_{B\Phi 3} + \delta D_{TP} + \delta D_{IOH} + \delta D_F + \delta D_{LM} + \delta D_{AMII} + \delta D_{PEJ} + \delta D_{SI}, \quad (14)$$

where  $X_{KA}(t_{2i})$  are coordinates of a SC position as of the time of radiation of signal  $t_{2i}$ ;

$X_{MII}(t_{3i})$  are meter vector coordinates as of the time of receiving signal  $t_{3i}$ ;

$D_{2i}(X_{KA}(t_{2i}), X_{MII}(t_{3i}))$  is the geometric value of a range to the SC;

$\delta D_{B\Phi 3}$  is the correction for phase delays of vehicle-borne equipment;

$\delta D_{TP}$  is the correction for tropospheric refraction;

$\delta D_{IOH}$  is the correction for ionospheric refraction;

$\delta D_F$  is the correction for discrepancy of phases and frequencies of SC generators and the meter;

$\delta D_{LM}$  is the correction for shift of the antenna phase centre with reference to SC centre of mass;

$\delta D_{AMII}$  is the correction for shift of the meter antenna phase centre;

$\delta D_{PEJ}$  is the relativistic correction;

$\delta D_{SI}$  is the correction for an instrumentation error.

Selection of a particular mathematical model for imitation of natural factors when you calculate the right side of the equation (14) is determined by requirements to precision of necessary parameters definition. In case of navigation non-requesting measurements the following parameters should be defined: orbital parameters; clock parameters and models

parameters. Orbital parameters include six orbital units, three light pressure scaling factors and empiric speed-up factors, and among the latter ones cyclical factors are the most important. Clock parameters include: current parameters of the clock on board SC and non-requesting measuring stations (NMS). Models parameters include corrections for: radio waves propagation delays in ionosphere and troposphere; non-uniqueness of phase modifications; position of SC centre of mass with reference to phase centres of SC and NMS antennas, instrumental noise, etc. In terms of this chapter's subject analysis of the task of clock current parameters simulation is of the utmost interest.

## 6.2 Models of instability of quantum frequency standards

According to the classical concept of clock rate instability, there are long-term and short-term components of its deviation from the uniform time scale. For example, in the book by (Tryon, 1983) a system of linear stochastic difference equations, describing the process of atomic time deviation in time domain  $S(k)$  and frequency domain  $q(k)$ , as well as of frequency drift  $W(k)$  is given:

$$\begin{aligned} S(k+1) &= S(k) + q(k)h + 0.5w(k)h^2 + V_s(k), \\ q(k+1) &= q(k) + w(k)h + V_q(k), \\ w(k+1) &= w(k) + V_w(k), \end{aligned} \quad (15)$$

Where  $V_s(k)$ ,  $V_q(k)$ ,  $V_w(k)$  are centred Gaussian processes by type of white noise with spread characteristics:  $\sigma_s, \sigma_q, \sigma_w$ ;  $h = t_{k+1} - t_k$  discretization interval of processes.

The given stochastic equations include both regular long-term and short-term stochastic components of atomic clock instabilities. They are used in recurrent procedures by Kalman type for estimation of amounts  $S(t)$ ,  $q(t)$ ,  $w(t)$  in tasks of time scales formation of group keepers, in synchronization tasks and so on.

In the book by (Oduan & Gino, 2002) a mathematical model of quantum clock is reviewed, which describes the relation of the noise power spectral density at different frequencies by type of:

$$S_y(f) = \sum_{\alpha=-2}^2 h_\alpha f^\alpha, \quad (16)$$

where  $h_\alpha$  are factors that determine power density of some noise components with frequencies  $f^\alpha$ .

Depending on integral number value  $\alpha$  it is assumed that there are 5 types of noise processes. For example, when  $\alpha$  changes from -2 to +2 with interval equal to 1, formula (16) describes respectively: white phase noise; flicker phase noise; white frequency noise and frequency random walk noise. Mean square two-sample dispersion (Alan variations) is connected to noise power  $S_y(f)$  by the relation (Oduan & Gino, 2002):

$$\sigma_y^2(\tau) = \int_0^\infty |H_A(f)|^2 |H_f(f)|^2 S_y(f) df, \quad (17)$$

where  $|H_A(f)|^2 = 2 \frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2}$  is a transfer function square modulus of a frequency digital filter in divergent integral (17).

$|H_f(f)|^2 = \begin{cases} 1 & \text{if } f < f_h \\ 0 & \text{if } f \geq f_h \end{cases}$  is a transfer function square modulus of a low-frequency filter with cutoff frequency  $f_h$ .

Formula for Alan dispersion computation is the result of integration of formula (17) with account of formula (16) and in view of condition:  $2\pi f_h \tau \gg 1$ ,

$$\sigma_y^2(\tau) = \frac{3h_2 f_h}{4\pi^2 \tau^2} + \frac{h_1}{4\pi^2 \tau^2} [1,04 + 3\ln(2\pi f_h \tau)] + \frac{h_0}{2\tau} + 2h_{-1} \ln 2 + \frac{2}{3} \pi^2 h_{-2} \tau. \quad (18)$$

Factors  $h_\alpha$  included into (18) are determined by preset Alan variations' values  $\sigma_y^2(\tau_i)$  in the left side of equation. In order to receive a unique definition of all five factors  $h_\alpha$  in the right side of equation (18), it is necessary to set up a system of five equations. Five values of Alan variations are determined by statistical processing of results of atomic clock rate measurements during five intervals  $\tau_i$  of different sizes. Normally intervals of 1, 10, 100 seconds, 1 hour and 1 day are chosen. As a result of solution of the system of equations set up in this way, five values of factors  $h_\alpha$  are determined. By inserting the determined values  $h_\alpha$  into equation (18), we receive a model of an observed clock instability, which is further used to make forecast of rate of scale deviation from a nominal value. However, the given classical model is not applicable to description of instabilities of some types of quantum clock. In particular, in terms of interval characteristics of instabilities typical for rubidium frequency standards, solution of the system of equations composed of (18) includes negative values of some factors  $h_\alpha$ , which is theoretically impossible. This model works only in case of clock, for which white frequency noise is predominant. The reason of restrictions on application of the classical model to description of instabilities of some kinds of quantum clock is that it does not determine extent of influence of some noise components on actual characteristics of instabilities of an observed clock.

A model of instability of quantum frequency standards (QFS) offered in researches by (Tissen, Tolstikov, 2004, 2011) enables to simulate by means of a computer a random process of QFS scale deviation from a nominal value for all probability distributions. It enables in its turn to check adequacy of application of a particular QFS physical model to describe deviation of the observed scale from the nominal value. In this context such a model can be regarded as software tool for metrological control of QFS physical models that are already worked out and just being developed. The algorithm development is based on the supposition of clock deviation from the nominal value in the form of a random recurrent process:

$$x_i = x_{i-1} + f_0 \tau + \delta x_i, \quad (19)$$

where  $x_i$  is clock rate variation as of the time point  $t_i$ ;  
 $f_0$  is generator rated frequency;  
 $\tau$  is discretization interval;  
 $\delta x_i$  is a random variable, determining root-mean-square error of clock deviation.

Variable  $\delta x_i$  in each interval is calculated according to the formula:

$$\delta x_i = \int_{t_{i-1}}^{t_i} Y_i dt, \tag{20}$$

where  $Y_i(t) = \sum_{j=1}^N y_j(t)$  is relative variation of generator frequency, composed of  $N$

components of frequency variations  $y_j = \frac{f_0 - f_j(t)}{f_0}$ ;

- $f_0, f_j(t)$  is an initial value and a current value of generator frequency;
- $i$  is number of time interval;
- $j$  is number of component of frequency variation.

In Figure 2 a block scheme which illustrates method of the described time scale formation algorithm in case  $N = 6$  is given.

A standard programmable random number generator (RNG) synthesizes 6 random number groups  $r_1, r_2, \dots, r_6$  with preset statistical characteristics of spread in 6 intervals of the following sizes:  $1, 10, \dots, 10^5$  sec. Time scale formation is the result of superposition of six random number groups generated by RNG. Accumulated groups of clock indications are analyzed according to the statistical Alan method to determine statistical characteristics of the time scale being simulated.

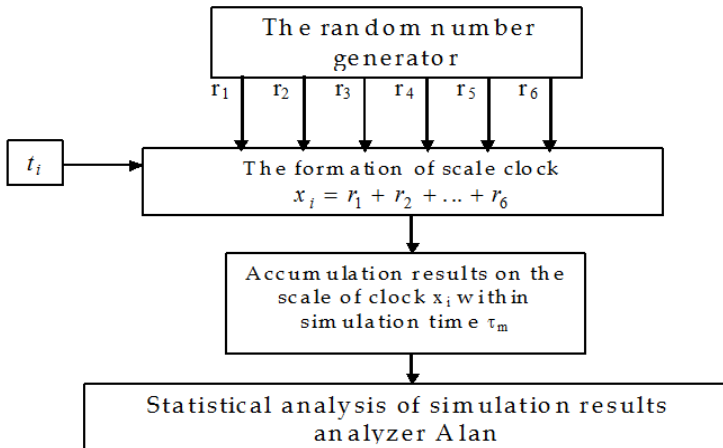


Fig. 2. Block scheme of quantum clock instability model

## 7. Comparison of ERP determination and prediction methods

### 7.1 General information

An almost 200 year long era of precise time determination with conventional astronomical methods ended in 2010. New satellite and radio interferometric techniques with about 2-3-

fold better measurements' resolution have substituted them. However, astrometric observations of stars, moon and planets positions accumulated through over the past centuries are invaluable for an overall picture of the Earth's rotation from the middle of the XVII century to the present. It is impossible to create adequate prediction models of Earth Rotation Parameters (ERP), such as universal time and pole coordinates, without use of these observations over a large enough time periods. A number of applied and fundamental Earth sciences have a need for ERP accurate data and predicted values for different periods. A study of fine structures in Earth rotation speed variations and its poles movements is the most urgent challenges of geodesy, astronomy and geodynamics.

According to a data cited in (Finkelshtein, 2007) the accuracy of ERP in IERS is about 50-60 microseconds of arc for pole coordinates and nutation angles, and about 4-10 microseconds for Universal Time. Achievement of such a high accuracy in comparison to classical methods, where the same estimation was usually made with a tolerance of approximately 1-2 ms, is now possible due to the comprehensive use of new means and techniques of the Measurement Assurance and Earth Rotation Parameters Prediction (MAERPP). The MAERPP Complex functionality includes the following measurement assurance technologies:

- Very Long Baseline Interferometers (VLBI);
- Global Navigation Satellite Systems (GNSS);
- Satellite Laser Ranging Systems (SLR).

Each of the methods for Earth rotation measurement has its own advantages, disadvantages and specific sources of systematic errors. Therefore, combining ERP series obtained by different methods and their combination in a single solution is an effective way to minimize systematic and random errors in ERP definitions.

## 7.2 Definition of Earth rotation parameters means VLBI

The VLBI is the most high-precision and independent ERP determination technique among the ones stated above, because stationary objects of the Universe named quasars are being observed. This method's principle of the geodynamics parameters determination consists in a measurement of the delay time of the same radio signal received by radio telescopes located at a distance from each other. Figure 3 shows the operation principle of the VLBI.

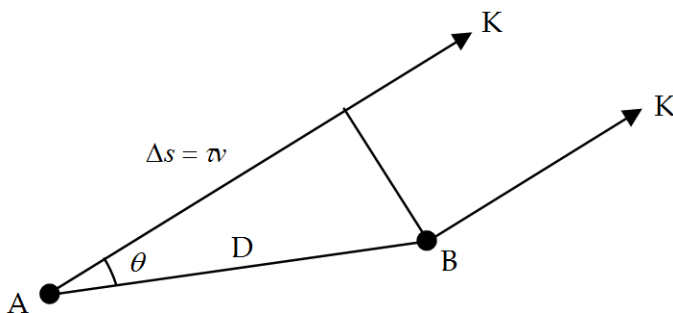


Fig. 3. The measurement principle of the VLBI.

The radio telescopes are installed in the points  $A$  and  $B$  at the distance  $D$  from each other; the interferometer's bases receive the radiation of the same quasar  $K$  in the centimetric spectrum synchronously. Processing of radio signals recorded allows determination of the time delay of  $\tau$  moments of radio wavefronts' arrival at the radio telescope  $A$  relative to the  $B$ , as well as interference frequency  $f$ . The difference between the distances from the radio telescopes to the quasar at the moment of observation  $\Delta s = \tau v$ , where  $v$  - radio propagation velocity. As a result of coprocessing of multiple values of  $\tau$  and  $f$  the following parameters can be determined: the difference in geocentric coordinates between the ends of the radio interferometer's base, which is used for determination of the  $D$  length and  $\theta$  angle, source's and the Earth's pole coordinates; the Earth's rotation instantaneous velocity; the elements of precession and nutation; the Greenwich apparent time of  $\tau$  and  $f$  determination moments; and other parameters.

### 7.3 Navigation and laser methods of the ERP determination

By the navigation and satellite laser ranging methods the ERP are being determined jointly with the spacecrafts' orbits parameters subsequent to the results of aggregate measurements from a large number of observation stations. The parameters, which have to be defined more precisely, may also include parameters of the models of atmosphere, Earth's gravity field, solar radiation, etc. The principles and methods of determination of the parameters named above are described quite full in a number of books (Urmaev, 1981), (Duboshin, 1983). However, the descriptions in literature are mostly general in nature. In order to implement them by means of software an additional work on calculation algorithms compilation has to be done first. The algorithm for joint determination of ERP and spacecrafts' orbits parameters, which can be used as a basis for ERP and spacecrafts' orbits parameters software compilation, is shown below.

If the earth referenced coordinates  $\bar{X}, \bar{Y}, \bar{Z}$  of a station tracking the passages of the spacecraft are known and the preliminary approximate values of the  $\bar{X}_p, \bar{Y}_p$  pole coordinates and the differences  $UT(t) = UT1 - UTC$  are known also, then the station position vector  $X, Y, Z$  at the time  $t$  in the middle equatorial coordinate system can be calculated by the following formula:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R_x(-\varepsilon) R_z(\Delta\Psi) R_x(\varepsilon + \Delta\varepsilon) R_z(-S_{\otimes}) R \begin{pmatrix} - \\ X_p, \bar{Y}_p \\ - \\ \bar{Z} \end{pmatrix} \begin{pmatrix} - \\ \bar{X} \\ - \\ \bar{Y} \\ - \\ \bar{Z} \end{pmatrix}, \quad (21)$$

where the ERP orientation matrixes are specified as usual:

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{pmatrix}, \quad R_z(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R \begin{pmatrix} - \\ \bar{X}_p, \bar{Y}_p \\ - \\ \bar{Z} \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\bar{X}_p \\ 0 & 1 & \bar{Y}_p \\ \bar{X}_p & -\bar{Y}_p & 1 \end{pmatrix};$$



$\varepsilon$  - is the obliquity of the movable equator plane to the plane of the instantaneous ecliptic;  
 $\Delta\psi$  - is the nutation in longitude;  
 $\Delta\varepsilon$  - is the nutation in obliquity;  
 $S_{\oplus}$  is the Greenwich apparent sidereal time calculated at the moment  $(t + \Delta UT(t))$ .

The calculated topocentric distance:

$$\rho_c(t) = \sqrt{(x-X)^2 + (y-Y)^2 + (z-Z)^2}, \quad (22)$$

where  $x, y, z$  are the coordinates of the spacecraft in the middle equatorial coordinate system.

Discrepancies  $\Delta\rho(t) = \rho_0(t) - \rho_c(t)$  are caused both by pseudorange measurement errors  $\rho_0(t) = c(t_{rec} - t_{rad})$  and by errors of an accepted computation model for the  $\rho_c(t)$  value, which depends on the spacecraft coordinates  $\Delta x, \Delta y, \Delta z$  errors and station coordinates errors  $\Delta X, \Delta Y, \Delta Z$ , as it follows from the equation (22).

Assuming that these values are small, the equation for discrepancies can be expanded in a Taylor series with the first term of the series only:

$$\Delta\rho_c = A \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} + B \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}, \quad (23)$$

where  $A$  and  $B$  vectors from the ground station to a satellite and back are as follows:

$$A = \begin{pmatrix} \frac{x-X}{\rho_c} & \frac{y-Y}{\rho_c} & \frac{z-Z}{\rho_c} \end{pmatrix}, \quad B = -A. \quad (24)$$

The transition from the formula (22) to the equation (23) is called linearization.

The spacecraft's coordinates  $x, y, z$  have a complex dependence on mean orbital elements in  $T_0$  epoch;  $\omega$  is the ascending node-perigee angle;  $n$  is the mean motion;  $\Omega$  is the longitude;  $i$  is the obliquity;  $e$  is the eccentricity;  $M$  is the mean anomaly.

The coordinates of the station depend on the position of the pole.  $X_p, Y_p$ .

Since the value  $\Delta UT$  belongs to the sidereal time calculation formula  $S_{\oplus}(t + \Delta UT(t))$ , the following substitute can be made:  $\Delta UT(t) \rightarrow \Delta S_{\oplus}(t)$  in rad.

Linear relations between the coordinates  $X, Y, Z$  and orbital elements of the spacecraft can be represented as a matrix:

$$\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = C \begin{pmatrix} \Delta\omega \\ \Delta\Omega \\ \Delta i \\ \Delta e \\ \Delta M \\ \Delta n \end{pmatrix}, \quad (25)$$

where  $C$  is the sensitivity matrix of the spacecraft coordinates to errors of the orbital elements:

$$C = \begin{pmatrix} \frac{\partial x}{\partial \omega} & \frac{\partial x}{\partial \Omega} & \frac{\partial x}{\partial i} & \frac{\partial x}{\partial e} & \frac{\partial x}{\partial M} & \frac{\partial x}{\partial n} \\ \frac{\partial y}{\partial \omega} & \frac{\partial y}{\partial \Omega} & \frac{\partial y}{\partial i} & \frac{\partial y}{\partial e} & \frac{\partial y}{\partial M} & \frac{\partial y}{\partial n} \\ \frac{\partial z}{\partial \omega} & \frac{\partial z}{\partial \Omega} & \frac{\partial z}{\partial i} & \frac{\partial z}{\partial e} & \frac{\partial z}{\partial M} & \frac{\partial z}{\partial n} \end{pmatrix}. \quad (26)$$

Linear relations for the station coordinates are as follows:

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = D \cdot \begin{pmatrix} \Delta x_p \\ \Delta y_p \\ S_{\oplus} \end{pmatrix}, \quad (27)$$

where  $D$  is the sensitivity matrix of the spacecraft coordinates to ERP errors:

$$D = \begin{pmatrix} \frac{\partial X}{\partial x_p} & \frac{\partial X}{\partial y_p} & \frac{\partial X}{\partial S_{\oplus}} \\ \frac{\partial Y}{\partial x_p} & \frac{\partial Y}{\partial y_p} & \frac{\partial Y}{\partial S_{\oplus}} \\ \frac{\partial Z}{\partial x_p} & \frac{\partial Z}{\partial y_p} & \frac{\partial Z}{\partial S_{\oplus}} \end{pmatrix}. \quad (28)$$

Since the partial derivatives of along the spacecraft's ascending node longitude  $\Omega$  and the sidereal time  $S_{\oplus}$  are calculated as follows:

$$\frac{\partial x}{\partial \Omega} = \frac{\partial X}{\partial S_{\oplus}} = -y; \quad \frac{\partial y}{\partial \Omega} = \frac{\partial Y}{\partial S_{\oplus}} = x; \quad \frac{\partial z}{\partial \Omega} = \frac{\partial Z}{\partial S_{\oplus}} = 0, \quad (29)$$

the last columns of the matrixes  $C$  and  $D$  are linearly dependent. Hence it is impossible to determinate  $\Delta\Omega$  and  $\Delta UT$  at the same time, and therefore the value  $\Delta UT(t)$  at small time intervals shall be approximated by the linear function:

$$\Delta UT(t) = \Delta UT(t_0) + \frac{DR(t-t_0)}{86400}, \quad (30)$$

where  $DR$  (sec.day) is variation length of the day.

Combining the equations (27) and (28) we get:

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = \begin{pmatrix} \frac{\partial X}{\partial x_p} & \frac{\partial X}{\partial y_p} & \frac{\partial X}{\partial \dot{S}_{\oplus}} \\ \frac{\partial Y}{\partial x_p} & \frac{\partial Y}{\partial y_p} & \frac{\partial Y}{\partial \dot{S}_{\oplus}} \\ \frac{\partial Z}{\partial x_p} & \frac{\partial Z}{\partial y_p} & \frac{\partial Z}{\partial \dot{S}_{\oplus}} \end{pmatrix} \cdot \begin{pmatrix} \Delta x_p \\ \Delta y_p \\ \dot{\Delta S}_{\oplus} \end{pmatrix} \quad (31)$$

where  $\dot{S}_{\oplus} = n_{\oplus} + \Delta \dot{S}_{\oplus}$ ;  $n_{\oplus} \approx 0,72921 \cdot 10^{-4}$  radian in sec - mean motion;  
 $\Delta \dot{S}_{\oplus}$  is the correction for the accepted value  $n_{\oplus}$ .

$$\bar{D} = \begin{pmatrix} \frac{\partial X}{\partial x_p} = -z \cos S_{\oplus} & \frac{\partial X}{\partial y_p} = -z \sin S_{\oplus} & \frac{\partial X}{\partial \dot{S}_{\oplus}} = -y(t-t_0) \\ \frac{\partial Y}{\partial x_p} = -z \sin S_{\oplus} & \frac{\partial Y}{\partial y_p} = z \cos S_{\oplus} & \frac{\partial Y}{\partial \dot{S}_{\oplus}} = x(t-t_0) \\ \frac{\partial Z}{\partial x_p} = \bar{X} & \frac{\partial Z}{\partial y_p} = -\bar{Y} & \frac{\partial Z}{\partial \dot{S}_{\oplus}} = 0 \end{pmatrix}. \quad (32)$$

Then the following formula will be true for calculation of the  $DR$  value:

$$DR = -\frac{2\pi}{n_{\oplus}} \cdot \frac{\Delta \dot{S}_{\oplus}}{n_{\oplus}}, \quad (33)$$

In the above expressions for the partial derivatives (32)  $t_0$  is a fixed time, at which the ERP is determined.

Combining the equations (23), (25) and (31) we get the following conditional equation of the least-squares method:

$$\rho_0(t) - \rho_c(t) = A \cdot C \begin{pmatrix} \Delta \omega \\ \Delta \Omega \\ \Delta i \\ \Delta e \\ \Delta M \\ \Delta n \end{pmatrix} + B \cdot \bar{D} \begin{pmatrix} \Delta x_p \\ \Delta y_p \\ \Delta \dot{S}_{\oplus} \end{pmatrix}, \quad (34)$$

in the matrixes  $A, B, C, \bar{D}$  all the partial derivatives are known, since they are calculated on the accepted initial values of mean orbital elements at epoch  $T_0$ :  $\omega, \Omega, i, e, M, n$ , Earth Rotation Parameters  $x_p, y_p, \Delta UT(t)$  and coordinates of observatories  $\bar{X}, \bar{Y}, \bar{Z}$ .

Corrections  $\Delta \omega, \Delta \Omega, \Delta i, \Delta e, \Delta M, \Delta n, \Delta x_p, \Delta y_p, \Delta \dot{S}_{\oplus}$  are calculated on a number of the spacecraft's constellation observations from several stations at different points of time.

$N$  conditional equations can be worked out on the aggregate data obtained. Discrepancies in the conditional equations can be minimized by the least-squares method by working out the  $M$  normal equations corresponding to the number of the parameters to be determined:

$$\sum_{i=1}^N \frac{\partial \rho_0(t)}{\partial P_k} \Delta \rho(t) = \sum_{j=1}^M \left( \sum_{i=1}^N \frac{\partial \rho_c(t)}{\partial P_k} \frac{\partial \rho_c(t)}{\partial P_j} \right) \Delta P_j \quad (35)$$

where  $k = 1, \dots, M$ .

Calculated from equation (35) corrections for  $\Delta P_j$ , add a parameter  $P_j$  to the initial values and the process is repeated again until the value of discrepancies get below a specified value  $\varepsilon$ .

The algorithm of the differential method for joint correction of the ERP series and spacecraft's orbits parameters is applicable both to navigation and to laser ranging measuring instruments. The main difference between the Satellite Laser Ranging (SLR) method and the GNSS technology is that SLR uses not a radio range of wave length, but an optical spectrum for measurements. This reduces atmospheric effects, eliminates uncertainty of the multiline radio signal propagation and provides a potentially higher resolution of measurements. In addition, the laser ranging of satellites allows calculating a change rate of RLS positions coordinates in the global velocity field. Precision geodetic systems are created on the basis of SLR location points, which monitor the geodetic satellites. WGS-84 is the most famous of these, which consists of dozens of points, located on the Earth's surface. The coordinates of these points are determined constantly, that allows to register tectonic processes. Information processing is performed by specialized centres, where it goes through operational communication channels, for example, the Internet. It implements the main advantage of the laser ranging in comparison to the radio positioning i.e. the possibility to determine a displacement of SLR points on the surface and in height during the geophysical researches. Such networks have a zero accuracy grade according to the geodetic classification. They are a basis for development of other levels of engineering networks integrated into the global system WGS-84. However, despite these advantages the SLR means are less influential in the ERP determination compared to the GNSS and VLBI techniques, so they inferior to them in mass and immediacy of the observational data supply.

In order to compare the quality of the different methods for ERP determination let us present data obtained from the bulletin of the Russian main metrology centre State service of time, frequency and IERS's bulletin. Table 2 shows the systematic deviations by the ERP determination from the IERS data, obtained in 2010 in the leading Russian processing centres according to the: Glonass/GPS; Satellite Laser and VLBI within the network of points of the domestic and world stations (Bulletin E-141-144, 2010).

Techniques	Analysis centre	dUT1 ( $\mu$ as)	X (0,00001")	Y (0,00001")
GNSS (GI/GPS)	SSTF	- 16	+55	-43
GNSS (GI/GPS)	IAA	+ 6	+31	+52
Satellite Laser	IAA	+ 14	+61	+9
Satellite Laser	IAC	-	+58	-19
VLBI - 24 h	SSTF	+ 16	-58	+36
VLBI - 24 h	IAA	+ 9	+53	+47

Table 2. Systematic deviations by the ERP determination using different techniques in 2010 according to the Russian processing centres.

The following abbreviations are used in the table 2:

SSTF - State service of time, frequency and the Earth rotation parameters determination;  
 IAA - Institute for applied astronomy of the Russian Academy of Sciences (RAS);  
 IAC - Information-analytical centre TSNIIMASH.

Table 3 shows root-mean-square errors (RMS) of the ERP determinations obtained in 2010 by IERS according to the result of GPS, VLBI and SLR measurements processing in the leading processing centres (Bulletin B 285, 2011). Since there are more than 30 processing centres in the IERS Bulletins, we will state only the RMS values spread limits for ERP determinations in these centres.

Techniques	UT1 ( $\mu$ as)	LOD ( $\mu$ as)	X (0.000001")	Y (0.000001")
VLBI - 24 h	8-15	-	30-100	70-150
VLBI -Intensive	14-17	-	-	-
Satellite Laser	-	15-30	100-160	80-200
GNSS (GPS)	-	5-30	13-45	13-40

Table 3. Estimated accuracies of the ERP determinations by using different techniques in 2010 according to the IERS data.

Analyzing the data in the Tables 2 and 3 we can conclude that the VLBI technique is the most accurate for the universal time parameter definition and is inferior to the GNSS in accuracy of the pole coordinates determination. The SLR method shows the results with a similar accuracy of the pole coordinates determinations in comparison with VLBI and is also inferior to the GNSS technique. However, the SLR and GNSS techniques have similar results by determination of the length of the day. It should also be noted that the data obtained from the Russian processing centres (see Table 2.) conform to an international standard for ERP determination accuracy level.

## 7.4 ERP prediction

### 7.4.1 Statement and urgency of the problem

Prediction of the Earth's rotation parameters is an important activity of the International Earth Rotation Service and national services of time and frequency in some progressive countries. The need of ERP prediction is dictated first of all by fast developing space technologies for civil and military areas. The ERP values, calculated on the basis of observations' processing, are related to the time frames when these observations were made. Therefore the ERP data at the current time can get known only on the basis of their precomputed values. The problem is that the precomputed (ephemeris) ERP values belong to the initial conditions by solution of differential equations of the GNSS spacecrafts motion. This is particularly important when it becomes necessary to transfer the GNSS spacecrafts in an autonomous mode. The ERP prediction errors, ultimately, limit the maximum achievable accuracy of the time and coordinate support (TCS), as well as element of exterior orientation (EEO), because all other causes of errors can be minimized by means of new developments and improvements of the measuring instruments, as well as enhancements of physical and mathematical theories of environmental factors' effect on the results of navigation measurements. Besides the handling of the EEO challenges the high-precision ERP predictions are also necessary for solving of the TCS problems of civil and military facilities, which are during a long time located in points, inaccessible for GNSS signals; they are also important for challenges of geophysics and meteorology by building of matching models of

the internal structure of the Earth and global climate change, as well as for geodynamics and geodesy challenges by the more precise definition of the Earth orientation parameters relative to the inertial sidereal reference system and coordinates determination of points relative to the absolute coordinate system.

Actual requirements of the GNSS EEO are based on the fact that the influence of ERP determination and prediction errors on the overall TCS error is insignificant. For example, pursuant to this condition, the universal time prediction error for the specified 15 days period of the GLONASS spacecraft operation in the autonomous mode must not exceed a value close to 1 ms. Unfortunately, despite the high level of modern scientific and technological developments in the field of the ERP prediction, an achievement of such level of accuracy is still a complicated and unsolved problem.

### 7.4.2 Proposed solutions

The proposed methods are based on the presentation of a number of known UT1 observations data at long periods of 100 years or more in form of a superposition of a large number of harmonics. Such an approach contradicts to some extent the currently prevailing point of view that non-periodic trend variations in the Earth's rotation speed influence much more than periodic components of the Earth's rotation irregularity. For example, according to data cited in (Belotserkovskiy, Kaufman, 1972), contribution to the overall universal time prediction error of 1 year due to the unpredictable trend variations is about 99% compared with the contribution due to the imperfection of tidal and seasonal models of the Earth's rotation speed variations. It allows to make an unequivocal conclusion that the only way to improve the reliability and accuracy of the ERP prediction is connected with a solving of the problem and a development of the prediction method most sensible to recent trend variations. According to the authors' investigations, the creation of such model is possible only subject to involvement of as much as possible known data for estimation of the prediction parameters.

### 7.4.3 Structure of the source data

Thanks to astronomical observations data of the moon and planets motion, maintained from the middle of the seventeenth century, we now have an opportunity to reconstruct the picture of the Earth's rotation over the past 350 years at least approximately. The Figure 4 represents a diagram of the  $UT1$  parameter variations starting from the 1656 to the present day in the Terrestrial Dynamic Time system  $TT$ . The diagram was drawn on the basis of the  $\Delta T = UT1 - TT$  differences data, represented in (AE SSSR, 1970 and Konstantinov A. I, Fler A.G., 1971).

Analyzing the diagram in the Figure 4 we can note that the Earth's rotation speed in past centuries varied slightly and rather evenly until the late nineteenth century, when a sharp slowdown occurred. During the past 100 years some periods of Earth's rotation deceleration and acceleration can be traced with specific intervals of 20-30 years. Nowadays the Earth's rotation slowdown is observed, which is likely to be continued during the next 15 to 20 years.

In the Figure 5 taken from (IERS) the dots indicate the Earth's pole wandering in 1996-2000., and the solid line indicate the trajectory of the mean pole from 1890 to 2000.

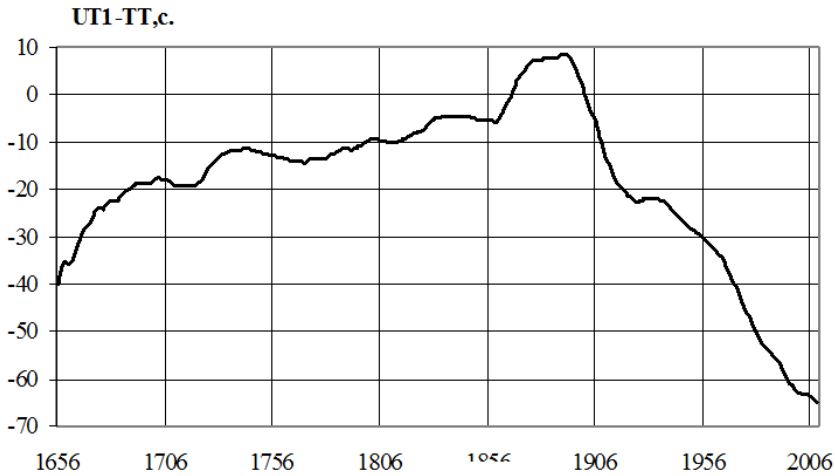


Fig. 4. Earth's rotation variations for the period from 1650 to 2010.

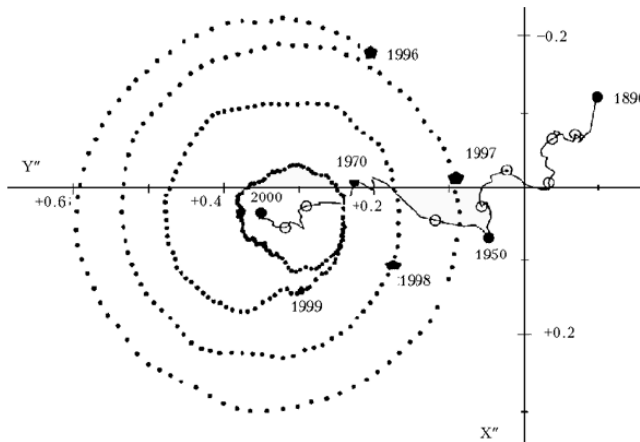


Fig. 5. Wandering of the Earth's North Pole

Analyzing the data, represented in the Figure 5, we can note that during the whole time of observations the mean pole was being displaced at a speed of about 10 cm/year along a complicated zigzag curve mainly in the direction of North America (meridian 290 ° EL).

**7.5 Overview of the proposed ERP prediction method**

The proposed method is based on approximation of periodic and quasi-periodic processes of the Earth's rotation using a polyharmonic mathematical model built on an ERP observations sample for at least 100 years. The calculation technique is based on the sequential adjustment of the original ERP time series by corrections introduction for tidal harmonics (for  $dUT1$  series), calculated on the basis of known physical and mathematical theory of the sun and moon motions, as well as harmonics corrections, obtained with

statistical methods (for all ERP series) . Generalized formula for  $dUT1$  parameter prediction can be written down as follows:

$$dUT1_{prg,i} = dUT1_{real,0} - dUT1_{pa,0} + dUT_{pa,i} + \Delta T_{pr,i} + \Delta T_{tr,i} + \Delta T_{sz,i}, \quad (36)$$

where  $dUT1_{prg,i}$  is the prediction of  $dUT1$  for the  $i$  -day;  
 $dUT1_{real,0}$  is the known value of  $dUT1$  for the reference day;  
 $dUT1_{pa,0}$  is the prediction of  $dUT1$  for the reference day made by means of the autoregression;  
 $dUT1_{pa,i}$  is the prediction of  $dUT1$  for the  $i$ -day made by means of the autoregression;  
 $\Delta T_{pr,i}$ ,  $\Delta T_{tr,i}$ ,  $\Delta T_{sz,i}$  are the corrections for the lunar and solar tides in the oceans, and for trend and seasonal variations at the  $i$  -day.

Tidal harmonics corrections are being calculated in accordance with the method accepted in the IERS (McCarthy, 2003). Consideration of the trend influence corrections by means of the harmonic model composed of harmonics with periods from 1 year to 70 years. Calculation of these harmonics parameters is performed by means of a specially developed for this purpose method of step-by-step summation of intervals of the time series. (Tissen et al., 2009). In order to calculate corrections for seasonal variations of the  $dUT1$  parameter the polyharmonic model shall be used as a sum of a year, half-year wave and an unlimited number of harmonics with periods less than a year. Calculation these harmonics parameters is performed at time intervals of about 15-25 years. The existing methods of  $dUT1$  prediction approximate these variations by a sum of year and half-year harmonics, whose parameters are estimated within the period of 4-6 years. In order to consider short-term stochastic components of the ERP series variations a mathematical model's component in a form of the autoregression shall be used as a restriction to smoothness.

Predictions for the pole coordinates shall be made in a similar way, except that in the equation (36) the influence of the lunar-solar tides shouldn't be taken into account because of their smallness.

## 7.6 Main results

Reports on the results of the ERP predictions obtained using the method developed in SNIIM from 2007 to 2011 were made many times at national conferences in Moscow and St. Petersburg, and at an international conference in Warsaw (Tissen & Tolstikov, 2011) in October 2009 (Tissen et al. , 2009).

Since October 2010 the method is being testing on the basis of the results of SNIIM/SSGA participation in the Earth Orientation Parameters Combination of Prediction Pilot Project (EOPCPPP). Participation in this project presupposes the daily data transfer into the main IERS centres: the Space Research Center of the Polish Academy of Science, Warsaw (eopcphp@cbk.waw.pl) and the United States Naval Observatory (USNO), Washington, District Columbia (eopcphp@maia.usno.navy.mil). Figures 5-7 show evaluations of the RMS predictions for universal time  $dUT1$  and pole coordinates:  $x_p$ ,  $y_p$  at intervals from 1 to 90 days by all EOPCPPP participants. The evaluations above, except for 3 participants, were received during the period of the project from September 2010 to November 2011



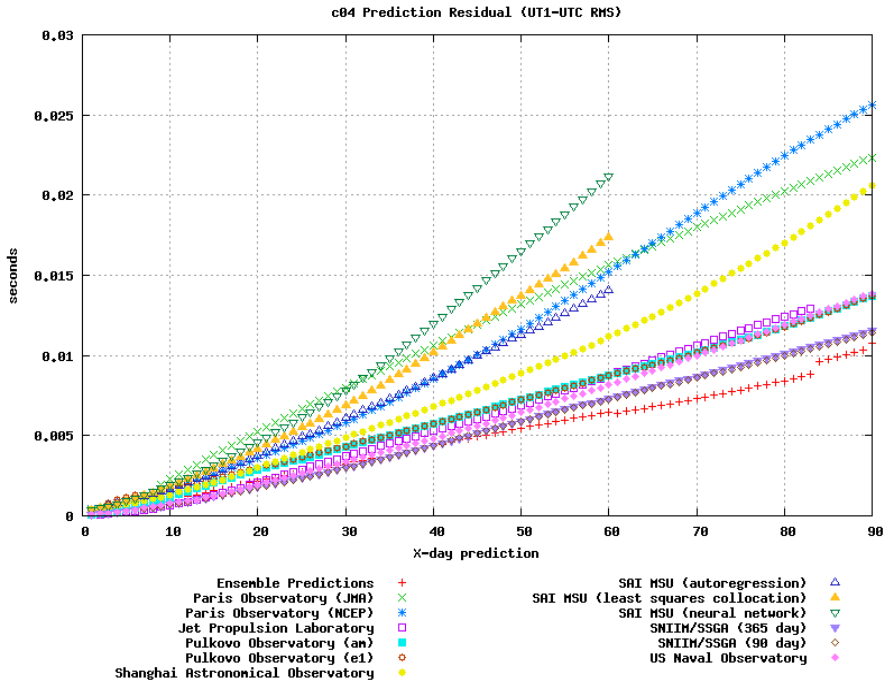


Fig. 6. Comparison of RMS for universal time predictions for 90 days made by EOPCPPP participants from September 2010 to November 2011.

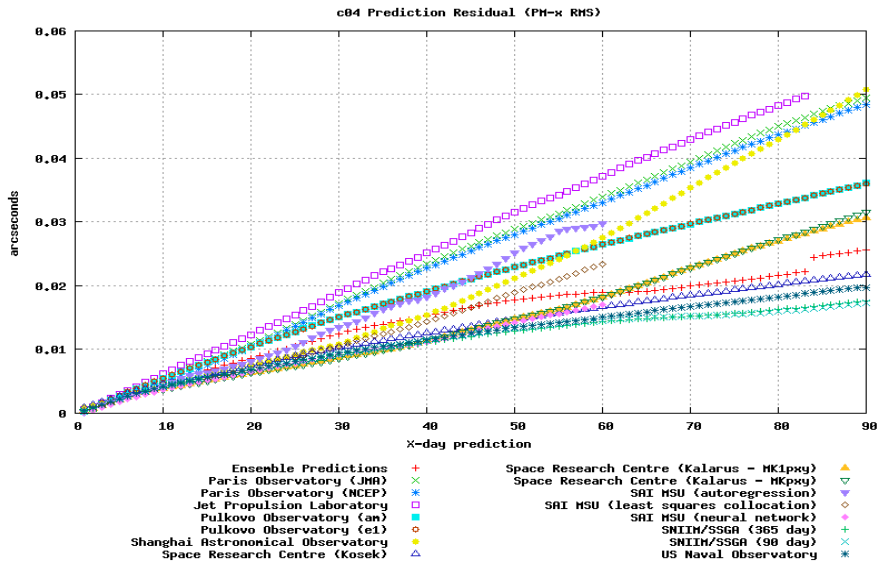


Fig. 7. Comparison of RMS for pole coordinates predictions:  $X_p$  and  $Y_p$  (top down, respectively) for 90 days made by EOPCPPP participants from September 2010 to November 2011.

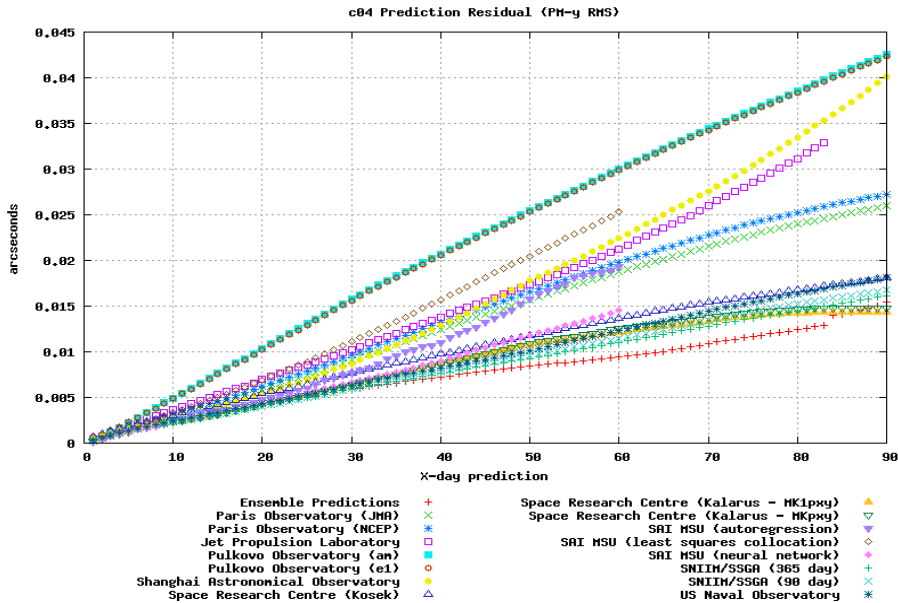


Fig. 8. Comparison of RMS for pole coordinates predictions:  $X_p$  and  $Y_p$  (top down, respectively) for 90 days made by EOPCPPP participants from September 2010 to November 2011.

Graphic data in the Figures 6-8 demonstrates an advantage of the ERP prediction method of the SNIIM/SSGA in comparison to the methods used by the other EOPCPPP participants, including those with official global suppliers of the ERP predictions used by the IERS (USNO).

During the first 13 months of the pilot project EOPCPPP, which involves more than 10 different methods for ERP prediction, the method used by SNIIM-SSGA showed the best results. In particular, a prediction accuracy of the universal time parameter, which is considered as the most important and difficult to predict one, within the period of less than 30 days was 1.19 times higher by the SNIIM - SSGA in 350 realizations than by the IERS and 1.93 times higher than by SSTF of Russia. Similar results were also obtained in longer intervals of up to 90 days and more. On the IERS website «eopcPPP@maia.usno.navy.mil» ERP predictions files from all the EOPCPPP project participants are published daily with their metrological characteristics according to the evaluations of RMS, absolute error (MAE) and standard deviation.

### 8. Conclusions

The issues stated in this chapter and the analyses were made on the basis of materials published in an open literature, as well as the results obtained directly by the authors. The historical information about the development of techniques and methods of time measuring, description of functions and responsibilities of the services of time and description of their workflow by example of the Novosibirsk time and frequency service are given for the familiarization purpose only. This information is necessary for

understanding of the most urgent issues and challenges of time, coordinate and navigational determinations, which successful solution depends largely on the time services' development level.

The simulation model of the quantum clock instability introduced in the fourth section is now used within the scope of the software and mathematical support of GLONASS. It can also be useful for verification of the adequacy of the metrological characteristics of existing and newly developed quantum frequency standards to accepted physical theories and models of their instability. The algorithms from the fifth section for codetermination of the spacecrafts' orbits parameters and ERP on the basis of ground (laser or navigation) measurements introduces the differential method for calculating of the parameters to be determined and can be used as a basis for development of the relevant software.

The sixth section represents comparative assessments of the ERP predictions made by the EOPCPPP project participants, according to which the quality of the SNIIM / SSGA predictions for the period from October 2010 to November 2011 on the majority of statistical evaluations for all ERP was preferable to all others, including the quality of the predictions provided by the USNO observatory. The good results obtained by the SNIIM/SSGA in the EOPCPPP project let us conclude that the right approach was used by the developing of the method. Particularly, the numerical analysis of ERP series of hundred years and more made by us shows that general regular compositions with specific periods of 2.4, 3.6, 4.8, 6.0 multiples of the main Chandler wobbles' period can be observed by the Earth's rotation variations and its poles wanderings. The exposed cross-correlation of the ERP series can be explained by the fact that the Earth's mass redistribution relative to its rotation axis occurs by the pole wandering. Consequently, the total moment of inertia relative to the Earth's rotation axis is changing leading to its angular velocity variations provided that the total angular momentum maintains. The discovered regularities and correlations lead us to the conclusion on the feasibility of creation of a generalized predictive model for the Earth's rotation, which combines all the Earth's angular velocity vector's components into a single consistent calculation pattern.

## 9. References

- Bakulin, P. & Blinov N. (1977). *Exact time service.*, pp. 320. Moscow: Science.
- Belotserkovskiy, D & Kaufman, M. (1972) *Assessment of the accuracy computing differences between the world and coordinated time with predictability up to a year. Research in the field of measurements of time and frequency/ /Proc. VNIIFTRI, Vol. 35*
- Duboshin, G. (1983) *Celestial mechanics. Methods of the theory of artificial celestial bodies* [text]/G. Duboshin. , pp. 650. Moscow: Science
- Great Russian encyclopedic dictionary. (2003). Moscow: The great Russian encyclopedia, pp. 1800.
- Finkelshtein, A. (2007). The fundamental coordinate-temporary ensuring. Bulletin of the Russian Academy of Sciences, p. 608-617, vol. 77.
- Hall, J. (2006) *Progress of natural sciences. Defining and measuring optical frequencies: prospects for optical clock – and not only.* Vol. 176, № 12.

- Ilyasov, Yu.; Korotkova, N. & Pshirkov, M. (2010). *Ensembles of pulsars for group timeline*. MGU
- Konstantinov, A; Fleer, A (1971). *Time*, pp. 368, Moscow: Publishing House standards.
- Kotyashkin, S. (1991) *Identification of the ionospheric delay signals in consumer equipment-frequency satellite NAVSTAR navigation system* // Overseas radioelectronics, № 1, p. 85-95.
- McCarthy, D. (2003). *Earth Orientation Parameters Prediction Research* / D. D. McCarthy, G. Petit // IERS Conventions (2003). – Frankfurt am Main : Verlag des Bundesamts für Kartographie und Geodäsie, 2004. p. 144-148, IERS Technical Note No. 32.
- Michal, S. (1983). *Clock. From gnomon to atomic clocks*. Moscow: Knowledge
- Oduan, K. & Gino, B. (2002). *A time dimension. GPS basics*/translated from Domnina, Y/edited by Tatarenkova, V. pp.400, Moscow: Tehnosfera
- Pipunyro, V. (1982). *History hours from ancient times to the present day*. pp. 496, Moscow: Science.
- Ryzhkov, A.; Novozhilov, E.; Legotin, N.; Koltunov, M. & Eriomin, E. (2007). *The ground segment of public service of time and frequency*. p. 42-44. //telecommunications. No. 2.
- Saastamoinen, J. (1972). *Inroduction to practical computation of astronomical refraction*. *Rull Jeadesigue*, 106, pp. 383-397.
- Storage, In: *Universal time, and the coordinates of the pole*, Bulletin E-141-144, (2010), Available from FGUP "VNIIFTRI", ISSN 0234-1069
- Storage, In: *International Earth Rotation and Reference Systems Service*, Bulletin B 285, 1.11.2011, Available from <http://www.iers.org/IERS/EN/Organization/organization.html>
- Tissen, V. & Tolstikov A. (2004). VII International conference " Actual problems of electronic instrument making -2004 ". *Mathematical model of instability of Quantum frequency standards [QFS]* p. 263-269. //Novosibirsk: NGTU. Volume 3.
- Tissen, V.; Tolstikov, A.; Balahnenko, A. & Malkin Z. (2009). Precision prediction of universal time in 100 years of data. // *Measuring equipment*. - 2009. № 12., 3-6.
- Tissen, V.; Tolstikov, A. & Malkin, Z. (2009). *UT1 prediction based on long-time series analysis*. In: IERS Workshop on EOP Combination and Prediction, Warsaw, Poland, 19-21 Oct 2009, Book of abstracts, p. 35.
- Tissen, V & Tolstikov, A (2011). Fourth all-Russian Conference. Fundamental and applied a temporary coordinate-and navigation [KVNO-2011]. *Universal model of quantum instability hours*. St. Petersburg, 10-14 October 2011.
- Tolstikov, A. (2011). Abstract methods and algorithms of coordinate-time definitions based on the use of satellite navigation technologies. SNIIM.
- Tryon, P. (1983). *Estimation of Parameters in Models cesium Beam atomic Clocs*. "J. Res. Nat. Bur. Stand.", 88, № 1, p. 3-16.
- Urmaev, M. (1981). *Orbital space geodesy methods*. p. 256. Moscow: Nedra.
- Unoshev, L. (1983). About comparisons of time signals navigation satellites. p. 30-33, *Measuring equipment*, № 8.

USSR Astronomy Yearbook (AE SSSR), 1970 pp. 672, Leningrad: Publishing House "Science"

# Relativistic Metrology: From Earth to Astrophysics

Luca Lusanna

*Sezione INFN di Firenze, Polo Scientifico, Sesto Fiorentino (FI)  
Italy*

## 1. Introduction

Till twenty years ago the basic concepts of special and general relativity were introduced by speaking of *clocks* and *rods* with an unspecified structure. Then the advances in atomic physics and in space navigation led to a revolution in metrology with the elimination of these old idealized notions and their replacements with realistic standards. For instance one can compare the 1965 point of view of Ref. (Basri S.A., 1965) with the 1997 one of Ref. (Guinot B., 1997).

However also today many scientists still think in terms of the *absolute notions of time and space* present in the Galilei space-time used in Newtonian physics, due to the fact that on Earth non-relativistic quantum mechanics is able to treat consistently problems ranging from molecular physics till quantum information without taking into account gravity (when needed Newtonian gravity is used). Only the description of light in atomic physics requires relativity (the trajectories of photons do not exist in Galilei space-time). Therefore most of the problems are formulated in inertial frames centered on inertial observers (having a constant velocity) in Galilei space-time (they are connected by the group of Galilei transformations containing space and time translations, spatial rotations and boosts) and, if needed, extended to non-relativistic accelerated frames taking into account the associated inertial apparent forces. The rotation of the Earth and its motion around the Sun are negligible effects for this type of physics.

Instead particle physics must face high speed objects and needs the Minkowski space-time of special relativity. Now the notions of space and time are no longer absolute: only the *global space-time is an absolute notion*. The Large Hadron Collider LHC particle accelerator at CERN is described with the coordinate time and the coordinate position of an inertial frame of Minkowski space-time centered on an inertial observer. To get the description with respect to another inertial observer one needs the group of Poincaré transformations (space and time translations plus Lorentz transformations  $\Lambda$ , i.e. spatial rotations and boosts;  $x'^{\mu} = a^{\mu} + \Lambda^{\mu}_{\nu} x^{\nu}$ ). In the new inertial frame the new coordinate time (and also the coordinate position) depends on both the old coordinate time and positions. This has generated an endless (and still going on) philosophical discussion on the meaning of time. Since the Lorentz-scalar line element joining two nearby points in an inertial frame of Minkowski space-time is  $ds^2 = (dx^0)^2 - \sum_r (dx^r)^2$  (with the particle physics conventions;  $ds^2 = -(dx^0)^2 + \sum_r (dx^r)^2$  with the general relativity ones;  $x^{\mu}$  are inertial time,  $x^0 = ct$ , and space,  $x^r$ , coordinates), space and time increments have different sign (Lorentz signature). The

only intrinsic structure of Minkowski space-time is that in each point A there is a light-cone (or null cone) defined by  $ds^2 = 0$ , which is the locus of the ray of light (traveling with the velocity of light) arriving in that point from the past or emanating from that point towards the future. The points inside the light-cone in A have a time-like distance from A and can be reached (if in the future) by traveling with a velocity less than the velocity of light. The points outside the light-cone of A have space-like distance from A and could be reached only with super-luminal velocity. If there is an atomic clock in A moving along a time-like curve towards the future its Lorentz-scalar *proper time* is defined as  $d\tau^2 = ds^2$  ( $d\tau^2 = -ds^2$  with the other convention) and coincides with the coordinate time  $x^0$  only in the inertial rest frame of the clock. However, since time is not absolute, there is no intrinsic notion of 3-space and of synchronization of clocks: both of them have to be defined with some convention. As a consequence the *1-way velocity of light* from one observer A to an observer B has a meaning only after a choice of a convention for synchronizing the clock in A with the one in B. Therefore the crucial quantity in special relativity is the *2-way (or round trip) velocity of light*  $c$  involving only one clock: the observer A emits a ray of light which is reflected somewhere and then reabsorbed by A so that only the clock of A is implied in measuring the time of flight. It is this velocity which is isotropic and constant in special relativity.

In Minkowski space-time the Euclidean 3-spaces of the inertial frames centered on an inertial observer A are identified by means of Einstein convention for the synchronization of clocks: the inertial observer A sends a ray of light at  $x_i^0$  towards the (in general accelerated) observer B; the ray is reflected towards A at a point P of B world-line and then reabsorbed by A at  $x_f^0$ ; by convention P is synchronous with the mid-point between emission and absorption on A's world-line, i.e.  $x_p^0 = x_i^0 + \frac{1}{2}(x_f^0 - x_i^0) = \frac{1}{2}(x_i^0 + x_f^0)$ . This convention selects the Euclidean instantaneous 3-spaces  $x^0 = ct = \text{const.}$  of the inertial frames centered on A. Only in this case the one-way velocity of light between A and B coincides with the two-way one,  $c$ . However, if the observer A is accelerated, the convention breaks down and we need a theory of non-inertial frames in Minkowski space-time as the one developed in Ref.(Alba D. et al, 2010, 2007). In this theory the transition from an inertial to a non-inertial frame (with its relativistic inertial forces and its non-Euclidean 3-spaces) can be described as a gauge transformation connecting two different generalized conventions for clock synchronization: therefore physics does not change, only the appearances of phenomena change.

However the International Space Station ISS near the Earth and space navigation in the Solar System require general relativity (at least its Post-Newtonian approximation) to take into account the effects of the gravitational field which is missing in special relativity. Now also space-time is no longer an absolute notion but is dynamically determined by Einstein's equations. Einstein's space-times have Lorentz signature but the structure of the light-cones changes from a point to another one. However rays of light, moving along null geodesics, are assumed to have the same 2-way velocity of light  $c$  as in special relativity, being an eikonal approximation to Maxwell equations. The equivalence principle implies that global inertial frames cannot exist: only locally near a particle in free fall we can have a local inertial frame and a local special relativistic approximation. Again there is the problem of clock synchronization for the definition of the non-Euclidean 3-spaces: even if the space-time is dynamically determined by Einstein's equations, each solution can be presented in arbitrary systems of 4-coordinates, since this is the gauge freedom of general relativity (form invariance of Einstein's equations under general coordinate transformations).

Therefore the presentation (gauge choice) of a solution of Einstein's equations is nothing else than a *metrology choice of a standard of space-time*, i.e. a choice of the time and space 4-coordinates.

Moreover this choice is fundamental for the description of macroscopic matter (its energy-momentum tensor is the source term in Einstein's equations) at the experimental level: physicists, space engineers and astronomers use an *intrinsically coordinate-dependent* (i.e. dependent on the chosen conventions) description of the trajectory of every macroscopic body (from spacecrafts to satellites, planets, stars,....).

Around the Earth *GPS (Global Positioning System) is a space-time standard* (Ashby N., 2003), relying on the time and length standards on the Earth surface. There is an array of 24 satellites around the Earth each one with an atomic clock with an accuracy which for modern commercial devices is today less than 30 nanoseconds. The satellites are at an altitude of 20,000 Km and have a mean velocity of 14,000 km/hr. Special relativity implies that these clocks tick more slowly (about 7 microseconds per day) than clocks on Earth. But general relativity implies that they tick faster (about 48 microseconds per day), so that a satellite clock advances faster than a clock on ground by about 38 microseconds per day. If we forget general relativity the precision of GPS localization (less than 15 meters) is lost within two minutes.

While in non-inertial frames in Galilei and Minkowski space-times there is a good understanding of the apparent inertial forces, in general relativity the gravitational field is described by the 4-metric tensor  ${}^4g_{\mu\nu}(x)$  in an arbitrary 4-coordinate system centered on an arbitrary observer (the line element is now  $ds^2 = {}^4g_{\mu\nu}(x) dx^\mu dx^\nu$ ) and it is not clear how to introduce a distinction between gravitational and inertial effects. However this is possible at the Hamiltonian level for the globally hyperbolic, asymptotically Minkowskian space-times, where it is possible to define global 3+1 splittings of the space-time, namely a foliation with 3-spaces evolving in time. The study of the 4-metric tensor in this framework allows one to disentangle the two physical degrees of freedom (or *tidal* variables) of the gravitational field (the two polarizations of gravitational waves in the linearized theory) and the gauge (or *inertial* variables) degrees of freedom describing the arbitrariness in the choice of the 4-coordinates. As shown in Ref. (Lusanna L., 2011) among the inertial variables there is the so-called York time (the trace of the extrinsic curvature of the 3-space as a 3-sub-manifold of space-time): it describes the remnant of the special relativistic gauge freedom in clock synchronization in this class of general relativistic space-times.

For the physics in the Solar System one assumes that the relevant Einstein space-times are globally hyperbolic (namely admitting a global definition of time) and asymptotically flat (namely tending to Minkowsky space-time at spatial infinity) space-times containing  $N$  bodies (the Sun and the planets) treated as point-like objects carrying multipoles (spin, moment of inertia,... of the extended body). A Post-Newtonian approximation is used in solving Einstein's equations in harmonic gauges and the gravitational waves inside the Solar System are shown to be negligible.

We have spoken only about Einstein general relativity. See Ref. (Will C.M., 2006, 1993) for the status of alternative theories of gravity inside the Solar System and Refs. (Soffel M.H., 1989; Brumberg V.A., 1991; Damour T. et al, 1991, 1992, 1993, 1994) for a treatment going from geodesy to celestial mechanics.



In what follows there will be a sketch, with update bibliography, of relativistic metrology inside the Solar System. It includes

A) Space and time standards.

B) The conventions needed for the description of satellites around the Earth: it is done by means of NASA (USA National Aeronautics and Space Administration) coordinates (Moyer T.D., 2003) firstly in the International Terrestrial Reference System ITRS (with an associated frame ITRF fixed on the Earth surface; see Ref. (IERS, 2004) for the IERS2003 conventions of the International Earth Rotation and Reference System Service IERS) and then in the Geocentric Celestial Reference System GCRS (with an associated non-rotating frame GCRF centered on the Earth center; see Ref. (Soffel M.H. et al, 2003) for the International Astronomic Units IAU2000).

C) The conventions needed for the description of planets and other objects in the Solar System uses the Barycentric Celestial Reference System BCRS (with an associated quasi-inertial Minkowski frame BCRF, if perturbations from the Milky Way are ignored, centered in the barycenter of the Solar System and with the axes specified by means of fixed stars (quasars) in the Hypparcos catalog (Kovalesky J. et al, 1989; Sovers O.J. et al, 1998; Ma C. et al, 1998; Johnstone K.J. et al, 1999; Fey A. et al, 2009)) and ephemerides (see IAU2000 in Ref. (Soffel M.H. et al, 2003)).

Ref. (Kaplan G.H., 2005) contains all the relevant aspects of these conventions.

While ITRF is essentially realized as a non-relativistic non-inertial frame in Galilei space-time, BCRF is defined as a *quasi-inertial frame, non-rotating* with respect to some selected fixed stars, in Minkowski space-time with nearly-Euclidean Newton 3-spaces. The qualification *quasi-inertial* is introduced to take into account general relativity, where inertial frames exist only locally. It can also be considered as a Post-Minkowskian space-time with 3-spaces having a very small extrinsic curvature. GCRF is obtained from BCRF by taking into account Earth's rotation around the Sun with a suitable Lorentz boost with corrections from Post-Newtonian gravity. By taking into account the extension of the geoid and Earth revolution around its axis one goes from the quasi-Minkowskian GCRF to the quasi-Galilean ITRF.

New problems emerge by going outside the Solar System. In astronomy the positions of stars and galaxies are determined from the data (luminosity, light spectrum, angles) on the sky, i.e. on a 2-dimensional spherical surface around the Earth with the relations between it and the observatory on the Earth done with GPS.

Then one needs a description of stars and galaxies as living in a 4-dimensional nearly-Galilei space-time with the International Celestial Reference System ICRS (see Refs. (Kovalesky J. et al, 1989; Sovers O.J. et al, 1998; Ma C. et al, 1998; Johnstone K.J. et al, 1999; Fey A. et al, 2009)), whose materialization ICRF is considered as a "quasi-inertial frame" in a "quasi-Galilei space-time", in accord with the assumed validity of the cosmological and Copernican principles. Namely one assumes a homogeneous and isotropic cosmological Friedmann-Robertson - Walker solution of Einstein equations (the standard  $\Lambda$ CDM cosmological model). In it the constant intrinsic 3-curvature of instantaneous 3-spaces is nearly zero as implied by the CMB data (Bartelmann M., 2010; Bean R., 2009), so that Euclidean 3-spaces (and Newtonian gravity) can be used (all galactic dynamics is Newtonian gravity). See Ref. (Lindegren L. et al, 2003) for the IAU conventions for defining the astrometric radial

velocity of stars taking into account astrometric positions, spectroscopy of star light and light propagation in gravitational fields.

However, to reconcile all the data with this 4-dimensional reconstruction one must postulate the existence of dark matter and dark energy as the dominant components of the classical universe (Durrer R., 2011; Bonvin C. et al, 2011; Garret K. et al, 2011; Ross M., 2010) after the recombination 3-surface (a kind of Heisenberg cut between quantum cosmology and classical astrophysics)!

## 2. Standards of length and time

In this Section we discuss the existing standards for length and time.

### 2.1 Standard of length

In 1975 in the 15th Meeting CGPM of the General Conference on Weights and Measures (Meeting 15, 1975) the conventional value of the 2-way velocity of light was fixed to be  $c = 299792458 \text{ m s}^{-1}$ .

In 1983 the 17th Meeting CGPM of the General Conference on Weights and Measures (Meeting 17, 1983) adopted the following standard of length

*The meter is the length of the path traveled by light in vacuum during a time interval of  $1/c$  of a second.*

To measure the 3-distance between two objects in an inertial frame we send a ray of light from the first object, to which is associated an atomic clock, to the second one, where it is reflected and then reabsorbed by the first object. The measure of the flight time and the 2-way velocity of light determine the 3-distance between the objects.

This convention is compatible with the Euclidean 3-space of inertial frames in Minkowski space-time. When the technology will allow one to measure the deviations from Euclidean 3-space implied by Post-Newtonian gravity we will need a modified convention taking into account a general relativistic notion of length.

In astronomy the unit of length, defined in the IAU (1976) System of Astronomical Constants, is the *astronomical unit AU*, approximately equal to the mean Earth-Sun distance (Resolution 10, 1976; IBWM, 2006). It is the radius of an unperturbed circular Newtonian orbit about the Sun of a particle having infinitesimal mass, moving with an angular frequency of 0.017 202 098 95 radians per day. Measurements of the relative positions of planets in the Solar System are done by radar (or by telemetry from space probes): one measures the time taken for light to be reflected from an object using the conventional value of the velocity of light  $c$ .

Both for objects inside the Solar System and for the nearest stars one measure the distance with the *trigonometric parallax* by using the propagation of light and its velocity  $c$  in inertial frames. One measures the difference (the inclination angle) in the apparent position of an object viewed along two different lines of sight at two different times and then uses Euclidean geometry to evaluate the distance. The used unit in astronomy is the parsec, which is 3.26 light-years or  $3.26 \cdot 10^{16}$  meters.

However this convention cannot be used for more distant either galactic or extra-galactic objects (UCLA, 2007; Carrol B.W. et al, 2007). New notions like standard candles, dynamical parallax, spectroscopic parallax, luminosity distance,..... are needed. These notions involve

both aspects of light propagation in curved space-times and cosmological assumptions like the Hubble law (velocity-redshift linear relation). Therefore they belong to another type of metrology.

## 2.2 Atomic clocks and ACES

The time scales like the SI (International System of Units) *atomic second* are based on frequency standards for *microwave* atomic clocks based on isotopes like cesium ( $^{133}\text{Cs}$ ) and rubidium ( $^{87}\text{Rb}$ ) and with frequencies of the order of GigaHertz ( $10^9\text{ Hz}$ ). While the national standard agencies (National Institute of Standards and Technology NIST in USA, National Physical Laboratory NPL in United Kingdom, Paris Observatory in France, Physikalisch-Technische Bundesanstalt PTB in Germany, Istituto Nazionale di Ricerca Metrologica INRIM in Italy) maintain an accuracy of 1 nanosecond per day ( $1\text{ ns} = 10^{-9}\text{ s}$ ), many primary cesium atomic clocks using laser cooled atomic fountains have an inaccuracy less than 100 picoseconds per day ( $1\text{ ps} = 10^{-12}\text{ s}$ ) with the best ones approaching 10 ps per day (Bize S. et al, 2005; Parker T.E., 2010).

If atomic clocks operating on different quantum transitions are considered as ideal clocks in general relativity, then they measure the same proper time (and not a coordinate time) along their trajectory (Guinot B., 1997). See Ref. (Reynaud S. et al, 2009) and its bibliography for the experiments on the *universality of clock rates* (relative frequency ratios between different clocks are constant at a level of the order of  $10^{-16}$  per year). See also Ref. (Perlick V., 1987, 1994) for another general relativistic effect, the *second clock effect*, according to which two clocks synchronized at the same point, then separated and finally rejoined remain synchronized in Riemannian space-times like Einstein's ones but not in Weyl space-times.

A new family of *optical* atomic clocks in the region of  $10^{15}\text{ Hz}$  is developing quickly with the help of optical frequency-combs for direct optical frequency measurements. They allow one to reach a fractional frequency inaccuracy of better than  $10^{-17}$  (corresponding to better than 1 ps per day) (Gill P., 2005; Rosenbad T. et al, 2008; Ludlow A.D. et al, 2008; Chou C.W. et al 2010a) and will become relevant for metrology in the near future. Moreover optical clocks allow to verify the "time dilation effect" for relative speeds of less than 10 m/s or for a change in height near the Earth's surface of less than 1 meter (Chou C.W., 2010b).

In Ref. (Arias E.F., 2005) there is a review of time metrology with a comparison of various time scales, the use of GPS receiver for time transfer (see also Ref. (Petit G. et al, 2005)) and on the dissemination and access to the international time scales. See also Refs. (Lemondé P. et al, 2001; SIGRAV 2006) for the status of atomic clocks in space near the Earth or on spacecrafts inside the Solar System.

The Atomic Clock Ensemble in Space (ACES) mission of the European Space Agency ESA (ACES 2010; Cacciapuoti L. et al, 2007, 2008; Blanchet L. et al, 2000), to be launched in 2015, aims to put a new microwave atomic clock (PHARAO, Projet d'Horloge Atomique par Refroidissement d'Atome en Orbite) together with an active hydrogen maser (SHM, Space active Hydrogen Maser) on the International Space Station (ISS; height 400 Km, rotation period 90 min, inclination angle  $51.6^\circ$ ). The two clocks will generate an on-board timescale with an expected frequency instability and inaccuracy at the  $10^{-16}$  level. There will be a frequency comparison between the space clocks and ground clocks using microwave links: in particular ACES will give the first precision measurement of the gravitational redshift of the geoid, namely of the  $1/c^2$  deformation of Minkowski light-cone caused by the geo-potential.

### 2.3 Time scales

The fundamental conceptual time scale is the *SI atomic second* whose definition is (Resolution 1, 1956)

*The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom. This definition refers to a cesium atom at rest at a temperature of 0° K.*

It gives a precise and constant rate of time measurement for *observers local to the apparatus on the surface of the Earth, i.e. on the rotating geoid* (for them it is a unit of *proper time* (Guinot B., 1997)), in which such seconds are counted.

However from 1971 the conventional practical high-precision atomic time standard is a coordinate time (Guinot B., 1997), the *International Atomic Time TAI*. TAI is defined as a suitable weighted average of the SI time kept by over 200 atomic clocks (mainly cesium clocks) in about 70 national laboratories worldwide (Circular T263, 2009). The comparison of the clocks is done using GPS signals and two-way satellite time and frequency transfer.

The next step is to connect TAI to the time scales based the *Earth Rotation*, which were used in astronomical applications as telescope pointing, depended on the geographical location of the observer and were based on observing celestial bodies crossing the meridian every day. Two such scales are:

*Greenwich sidereal time* is the hour angle of the equinox measured with respect to the Greenwich meridian.

*Local sidereal time* is the local hour angle of the equinox or the Greenwich sidereal time plus the longitude (east positive) of the observer, expressed in time units. Sidereal time appears in two forms, *mean* (GAST Greenwich Apparent mean Sidereal Time or LMST Local Mean Sidereal Time) and *apparent* (LAST, Local Apparent Sidereal Time), depending on whether the mean or true equinox is the reference point. The position of the mean equinox is affected only by *precession* while the true equinox is affected by both *precession and nutation*. Let us remember that the equinox is a direction in space along the nodal line defined by the intersection of the ecliptic (the plane of the Earth's orbit) and equatorial planes. The difference between true and mean sidereal time is the *equation of the equinoxes*, which is a complex periodic function with a maximum amplitude of about 1sec. Of the two forms, apparent sidereal time is more relevant to actual observations, since it includes the effect of nutation. Greenwich (or local) apparent sidereal time can be operationally obtained from the right ascensions of celestial objects transiting the Greenwich (or local) meridian.

Nowadays *Universal Time UT* is the generic timescale based on Earth's rotation. It is determined by Very Long Baseline Interferometry (VLBI) observations of distant quasars with an accuracy of microseconds. There are various variants of UT. The most used is UT1, based on VLBI observations of quasars, on Lunar Laser Ranging (LLR), on determination of GPS satellite orbits. UT1 is the same everywhere on Earth and is proportional to the rotation angle of the Earth with respect to distant quasars.

An approximate version of UT1 is the *Coordinate Universal Time UTC*. It is an atomic time scale and the international standard for *civil time*. It is a hybrid time scale (ITUR, 2007), which uses SI atomic seconds on the geoid (it usually has 86 400 SI seconds per day), but subject to occasional 1 second adjustments (the so-called *leap second*) to keep it within 0.9 seconds from

UT1 ( $UT1 \approx UTC + DUT1$  with  $DUT1 \approx \pm 0.1 \text{ sec}$ ) and to have  $TAI = UTC + \Delta AT$  ( $\Delta AT$  is an integer number of leap seconds).

Other civil times given in Ref. (Moyer T.D., 2003) are

*GPS Master Time*: it is an atomic time for GPS receiving station on Earth and for GPS satellites  
-  $TAI = GPS + const.$ .

*ST Station Time*: it is an atomic time at a Deep Space Network (DSN) tracking station on Earth. It is assumed  $UTC \text{ or } GPS = ST + a + b(t - t_0) + c(t - t_0)^2$ .

As UT is slightly irregular in its rate, astronomers introduced Ephemeris Time and then replaced it with Terrestrial Time TT.

The *Ephemeris Time*  $T_{eph} = ET$ , replacing an old barycentric dynamical time TDB, is a relativistic coordinate time based on high-precision ephemerides, which are lists of instantaneous positions of the centers of mass of Sun, Moon and planets with respect to (equatorial rectangular 3-coordinates of) BCRS for any date and time between 1600 and 2001, developed at the Jet Propulsion Laboratory (JPL) and denoted DE405/LE405 (Kaplan G.H., 2005; Standish E.M., 1998). Lunar rotation angles are also provided. The DE405 coordinate system has been aligned to the ICRS. The JPL ephemerides are computed by an N-body numerical integration carried out in BCRS.

*Terrestrial Time TT*, which is an astronomical time scale used for geocentric and topo-centric ephemerides. The "standard epoch" for modern astrometric reference data, designated J2000.0 is expressed as a TT instant: J2000.0 means 2000 January 1, 12<sup>h</sup> TT at geo-center (Julian date JD 24515450 TT; J2000.0 is shorthand for the celestial reference system defined by the mean dynamical equator and equinox of J2000.0) (Kaplan G.H., 2005). TT is an idealized form of TAI ( $TT = TAI + 32^s.184$ ). TT runs at the same rate as a time scale based on the SI second on the surface of the Earth.

As shown in Eq.(2.6) of Ref. (Kaplan G.H., 2005) we have  $T_{eph} \approx TDB \approx TT + F(T)$ , where  $F(T)$  is a given function of the number T of Julian centuries of TT from J2000.0 ( $T = (JD(TT) - 2451545.0) / 36525$ ).

See also Ref.(Moyer T.D., 2003), p.18, where the following chains of transformations are defined

$T_{eph} \rightarrow TAI (\rightarrow UT1) \rightarrow UTC, GPS \rightarrow ST$  (ST = time scale of a tracking station on the Earth),

$T_{eph} \rightarrow TAI \rightarrow GPS \rightarrow ST$  (ST = time scale at an Earth satellite)

Let us remark that the astronomical universal time UT1 is defined by using the new earth precession-nutation theory denoted IAU2000A (relating the International Celestial Reference Frame ICRF to the International Terrestrial Reference Frame ITRF from 2003), which has been replaced in 2009 with a more dynamically consistent precession model denoted IAU2006 (Coppola V., 2009; IAU, 2006).

According to IAU2006 UT1 is *linear* in the Earth rotation angle  $\theta$ , a geocentric angle (such that  $\dot{\theta} = \omega_{earth}$  is the average angular velocity of rotation of the Earth) with a *Non-Rotating Origin NLO* in the equatorial plane orthogonal to the *Celestial Intermediate Pole CIP* from the axes centered in the *Celestial Intermediate Origin CIO* with no instantaneous rotation around

the Earth axis to the axes centered in the *Terrestrial Intermediate Origin TIO* rotating with the Earth (in IAU2000A CIO and TIO were called CEO, Celestial Ephemeris Origin, and TEO, Terrestrial Ephemeris Origin, respectively).

The lengths of the sidereal ( $\theta$ ) and UT1 seconds, and the value of  $\dot{\theta}$ , are not precisely constant when expressed in a uniform time scale such as TT. The accumulated difference in time measured by a clock keeping SI seconds on the geoid from that measured by the rotation of the Earth is  $\Delta T = TT - UT1$ . The long-term trend is for  $\Delta T$  to increase gradually because of the tidal deceleration of the Earth's rotation, which causes UT1 to lag increasingly behind TT. In predicting the precise times of topo-centric phenomena, like solar eclipse contacts, both TT and UT1 come into play, and this requires assumptions about the value of  $\Delta T$  at the time of the phenomenon. Alternatively, the circumstances of such phenomena can be expressed in terms of an *imaginary system of geographic meridians* that rotate uniformly about the Earth's axis ( $\Delta T$  is assumed zero, so that UT1 = TT) rather than with the real Earth; the real value of  $\Delta T$  then does not need to be known when the predictions are made. The zero-longitude meridian of the uniformly rotating system is called the *ephemeris meridian*.

Finally the astronomical conventions IUA2000 (Soffel M.H. et al, 2003) for the description of the Solar System (BCRS) and of the space near the Earth (GCRS) introduced the following two theoretical time scales not taken by any real clock but connected with Post-Newtonian solutions of Einstein's equations in special harmonic gauges with Sun, Earth, Moon, planets as matter.

*Barycentric Coordinate Time* -  $t_B = TCB$  - it advances at a rate  $1.55 \cdot 10^{-8}$  faster with respect to SI seconds on the surface of the Earth and is the time coordinate in BCRS.

*Geocentric Coordinate Time* -  $t_G = TCG$  - it advances at a rate  $6.97 \cdot 10^{-10}$  faster with respect to SI seconds on the surface of the Earth and is the time coordinate in GCRS. The connection to the terrestrial time is assumed to be  $TT = TCG - L_G (TCG - t_o)$  with a constant rate  $\frac{dT}{dt_G} = 1 - L_G$  with  $L_G = 6.969290134 \cdot 10^{-10}$ , while the transformation connecting TCB and TCG is given in the next Section.

Let us notice that the discussion whether it is better to use primary conventions based on atomic clocks or to revert to astronomical conventions is still open (Finkleman D. et al, 2011) and will be discussed again in 2012. For a recent update on the problem of time see Ref. (McCarthy D.D., 2009). At this stage it is difficult to say which point of view will become more relevant in the near future: how to compare astronomic precisions connected to VLBI and LLR with theoretical problems of atomic clocks like whether an atomic fountain clock can be approximated with a mass-point with a well defined proper time?

### 3. The space-time in the Solar System and near the Earth

In this Section we will describe the conventions used to describe physics on the Earth's surface and space physics near the Earth and in the Solar System. Instead of starting from the Earth, where Newtonian gravity is dominating, we shall begin with the general relativistic description of the Solar System.

The IAU conventions (Soffel M.H. et al, 2003) for the Solar System identify a system of harmonic coordinates (a BCRF frame) centered on the solar system barycenter and a Post-Newtonian solution of Einstein's equations in a special harmonic gauge at the  $O(1/c^3)$

order, which can be interpreted as an asymptotically-Minkowskian Post-Newtonian Einstein space-time. With a suitable coordinate transformation this solution is transformed in a description of the same space-time with new harmonic coordinates centered on the center of the Earth (a GCRF frame). However the presentation of this Einstein's space-time is strongly special relativistic (just  $O(1/c^2)$  for the NASA coordinates of spacecrafts (Moyer T.D., 2003)) and becomes Galilean when one makes the transition from the coordinates with origin in the center of the Earth to coordinates fixed on the crust of the Earth (a ITRF frame in the IERS conventions) (IERS, 2003).

### 3.1 BCRS - Barycentric Celestial Reference System

The resolution B1.3 of IAU2000 (Soffel M.H. et al, 2003) states that the Barycentric Celestial Reference System BCRS is a global reference system of barycentric space-time coordinates for the Solar System within the framework of general relativity. It is centered in the barycenter of the Solar System, which can be considered as a quasi-inertial Minkowski observer with a constant 4-velocity (the time axis of the barycentric time  $t_B = TCB$ ), because the effects of the Milky Way are negligible. Its spatial axes (in the instantaneous 3-spaces  $\Sigma_{t_B}$  with  $t_B = const.$  with rectangular 3-coordinates  $x_B^i$ ) are restricted to be *kinematically non-rotating*, namely they have no systematic rotation with respect to distant objects in the universe. For all practical applications the spatial axes are assumed to be oriented like the spatial axes of ICRS (see next Section). Therefore to each ICRF frame giving a materialization of ICRS is associated a BCRF frame.

The harmonic 4-coordinates and the retarded Post-Newtonian solution of Einstein's equations for the 4-metric  $g_{B\mu\nu}(x_B)$  given in the IAU2000 conventions are

$$x_B^\mu = \left( x_B^0 = c t_B; x_B^i \right),$$

$$g_{B00}(x_B) = \epsilon \left[ 1 - \frac{2w_B(x_B)}{c^2} - \frac{2w_B^2(x_B)}{c^4} + O(c^{-5}) \right],$$

$$g_{B0i}(x_B) = \epsilon \left[ \frac{4w_{Bi}(x_B)}{c^3} + O(c^{-5}) \right],$$

$$g_{Bij}(x_B) = -\epsilon^3 g_{Bij}(x_B) = -\epsilon \left[ \left( 1 + \frac{2w_B(x_B)}{c^2} \right) \delta_{ij} + O(c^{-4}) \right].$$

The signature of the 4-metric is the same as for Minkowski metric  $\eta^{\mu\nu} = \epsilon (+ - - -)$  ( $\epsilon = +$  is the particle physics convention,  $\epsilon = -$  is the general relativity one). The 3-metric  ${}^3g_{Bij}(x_B)$  on the 3-spaces  $\Sigma_{t_B}$  is positive-definite.

See Appendix A of Ref. (Soffel M.H. et al, 2003) for the Post-Newtonian gravitational potentials  $w_B(x_B)$  and  $w_B^i(x_B)$  generated by the Sun and the planets. These extended bodies are usually approximated with their center of mass (mass monopole) carrying, when needed like for Saturn, a spin dipole.

The barycenter of the Solar System has coordinates  $x_{B(B)}^\mu = (x_B^0; 0^i)$  and its world-line is a straight-line (the time axis) approximating a time-like geodesic of the 4-metric if we neglect galactic and extra-galactic influences. In each point of the barycentric world-line there is an *orthonormal tetrad* with the time-like 4-vector given by the barycenter 4-velocity  $u_{B(B)}^\mu = (1; \vec{0})$

and with the 3 mutually orthogonal spatial axes  $\epsilon_{B(B)r}^\mu = (0; \vec{\epsilon}_{B(B)r})$  whose orientation is determined by ICRS. The instantaneous 3-spaces  $\Sigma_{t_B}$  are considered as nearly Euclidean inertial 3-spaces. However their extrinsic 3-curvature as 3-sub-manifolds of the space-time is not zero but of order  $O(c^{-2})$ , so that strictly speaking they do not correspond to Einstein's clock synchronization convention.

The world-line  $x_{B(G)}^\mu(x_B^o) = (x_B^o; \vec{x}_{B(G)}(x_B^o))$  of the Earth's geo-center (a time-like geodesics of the Post-Newtonian 4-metric  $g_{B\mu\nu}(x_B)$ ) is determined by the *JPL ephemerides as solution of the equations of motion of the solar system bodies*. The geo-center has the 4-velocity  $u_{B(G)}^\mu(x_B^o) = \frac{d x_{B(G)}^\mu(x_B^o)}{d x_B^o} = (1; \vec{v}_{B(G)}(x_B^o))$  and carries spatial 3-axes  $\epsilon_{B(G)r}^\mu(x_B^o) = (0; \vec{\epsilon}_{B(G)r})$  assumed parallel to the axes  $\epsilon_{B(B)r}^\mu$  of the barycenter. For an arbitrary point in the solar system with coordinates  $(x_B^o; \vec{x}_B)$  we have  $\vec{x}_B = \vec{x}_{B(G)}(x_B^o) + \vec{r}_{B(G)}(x_B^o)$ .

The global reference system BCRS is the reference system in which the positions and motions of bodies outside the immediate environment of the Earth have to be expressed. It is appropriate for the solution of the equations of motion of solar system bodies (the development of the solar system ephemerides). Within it the positions and motions of galactic and extra-galactic objects are most simply expressed. It is the system to be used for most positional-astronomy reference data, e.g. star catalogues.

### 3.2 GCRS - Geocentric Celestial Reference System

The resolution B1.3 of IAU2000 (Soffel M.H. et al, 2003) states that the Geocentric Celestial Reference System BCRS is a global reference system of space-time coordinates for Earth based measurements and the solution of the equations of motion of bodies in the near-Earth environment (artificial satellites) within the framework of general relativity. The GCRS is defined such that the transformation between BCRS and GCRS spatial coordinates *contains no rotation component*, so that GCRS is kinematically non-rotating with respect to BCRS. The equations of motion of an Earth satellite with respect to GCRS will contain relativistic Coriolis forces that come mainly from geodesic precession. The spatial orientation of the GCRS is derived from that of BCRS, that is by the orientation of the ICRS. Its origin is the world-line of the geo-center (fictitious observer at the center of the Earth): it is the time axis of the geocentric time  $t_G = TCG$  and the instantaneous 3-spaces  $\Sigma_{t_G}$  with  $t_G = const.$  are inertial hyper-planes (Einstein's convention for clock synchronization) only at the lowest order in  $1/c$ .

GCRS has the following 4-coordinates and retarded Post-Newtonian solution of Einstein's equations for the 4-metric  $g_{G\mu\nu}(x_G)$

$$x_G^\mu = (x_G^o = c t_G; x_G^a),$$

$$g_{G00}(x_G) = \epsilon \left[ 1 - \frac{2w_G(x_G)}{c^2} - \frac{2w_G^2(x_G)}{c^4} + O(c^{-5}) \right],$$

$$g_{G0a}(x_G) = \epsilon \left[ \frac{4w_G^a(x_G)}{c^3} + O(c^{-5}) \right],$$

$$g_{Gab}(x_G) = -\epsilon^3 g_{Gab}(x_G) = -\epsilon \left[ \left( 1 + \frac{2w_G(x_G)}{c^2} \right) \delta_{ab} + O(c^{-4}) \right].$$



See Appendix A of Ref. (Soffel M.H. et al, 2003) for the GCRS Post-Newtonian gravitational potentials  $w_G(x_G)$  and  $w_G^a(x_G)$ . While  $w_G(x_G)$  generalizes the Newton potential, the components  $g_{G0a}(x_G)$  (i.e.  $w_G^a(x_G)$ ) are responsible for the *gravito-magnetic* effects near the Earth like the Lense-Thirring or frame-dragging effect (Ciufolini I. et al, 1995; Will C.M., 2011). The Post-Newtonian solution  $g_{G\mu\nu}(x_G)$  describes the *exterior* gravitational field outside the Earth surface, not inside.

The geo-center has coordinates  $x_{G(G)}^\mu = (x_G^0; 0^a)$  with tangent time-like vector (the unit 4-velocity)  $u_{G(G)}^\mu = (1; 0^a)$ , while the spatial axes have the 3 orthogonal tangent space-like unit vectors  $e_{G(G)r}^\mu$ . It is a time-like geodesic of the Post-Newtonian 4-metric  $g_{G\mu\nu}(x_G)$ , if the Earth is approximated as a *mass monopole*. Otherwise the Earth mass and spin multipoles will create a deviation of the geo-center world-line from a time-like geodesic.

The tetrad carried by the geo-center is obtained from the BCRS tetrad with the tensorial transformation law of 4-vectors, i.e. with the matrix  $\frac{\partial x_G^\mu}{\partial x_B^i} \Big|_{\text{geocenter}}$ . To evaluate it one needs the transformation between BCRS and GRCS coordinates. If the barycentric 3-velocity and 3-acceleration of the geo-center are  $\vec{v}_{B(G)} = \frac{d\vec{x}_{B(G)}}{dt_B}$  and  $\vec{a}_{B(G)} = \frac{d\vec{v}_{B(G)}}{dt_B}$  respectively and if we introduce the relative 3-vector  $\vec{r}_{B(G)}(x_B^0) = \vec{x}_B - \vec{x}_{B(G)}(x_B^0)$ , the BCRS-GCRS coordinate transformation is (Soffel M.H. et al, 2003)

$$\begin{aligned} t_G &= t_B - \frac{1}{c^2} \left[ A(t_B) + v_{B(G)}^i r_{B(G)}^i \right] + \\ &+ \frac{1}{c^4} \left[ B(t_B) + B^i(t_B) r_{B(G)}^i + B^{ij}(t_B) r_{B(G)}^i r_{B(G)}^j + C(t_B, \vec{x}_B) \right] + O(c^{-5}) = \\ &= t_B - \frac{1}{c^2} \left[ \int_{t_{B0}}^{t_B} dt \left( \frac{\vec{v}_{B(G)}^2}{2} + w_{Bext}(\vec{x}_{B(G)}) \right) + v_{B(G)}^i r_{B(G)}^i \right] + O(c^{-4}), \\ x_G^a &= \delta_{ai} \left[ r_{B(G)}^i + \frac{1}{c^2} \left( \frac{1}{2} v_{B(G)}^i v_{B(G)}^j r_{B(G)}^j + w_{Bext}(\vec{x}_{B(G)}) r_{B(G)}^i + \right. \right. \\ &\left. \left. + r_{B(G)}^i a_{B(G)}^j r_{B(G)}^j - \frac{1}{2} a_{B(G)}^i r_{B(G)}^2 \right) \right] + O(c^{-4}), \end{aligned}$$

The functions  $A(t_B)$ ,  $B(t_B)$ ,  $B^i(t_B)$ ,  $B^{ij}(t_B)$  and  $C(t_B, \vec{x}_{B(G)})$ , depending on the BCRS gravitational potentials  $w_{Bext}(\vec{x}_{B(G)})$  (the BCRS Newtonian potential of all solar system bodies apart from the Earth acting on the geo-center) and  $w_{Bext}^i(\vec{x}_{B(G)})$  (the BCRS gravito-magnetic potential) are given in Ref. (Soffel M.H. et al, 2003).

As shown in Ref. (Soffel M.H. et al, 2003), this transformation reduces to a pure Lorentz boost without rotation modulo terms of order  $O(c^{-4})$  in the limit of no acceleration due to the gravitational field (i.e. with  $\vec{x}_{B(G)}(t_B) = \vec{v}_{B(G)} t_B$ ,  $\vec{v}_{B(G)} = \text{const.}$ ,  $v_{B(G)} = |\vec{v}_{B(G)}|$ ,  $\beta_{B(G)} = v_{B(G)}/c$ ,  $\gamma_{B(G)} = (1 - \beta_{B(G)}^2)^{-1/2}$ )

$$\begin{aligned} t_G &= \gamma_{B(G)} \left( t_B - \frac{\vec{v}_{B(G)} \cdot \vec{x}_B}{c^2} \right) + O(c^{-4}), \\ \vec{x}_G &= \vec{x}_B - \gamma_{B(G)} \vec{v}_{B(G)} t_B + \frac{\gamma_{B(G)} - 1}{v_{B(G)}^2} \vec{v}_{B(G)} \cdot \vec{x}_B \vec{v}_{B(G)} + O(c^{-4}). \end{aligned}$$

Without the kinematically non-rotating constraint, GCRS would have a *slow rotation* ( $\approx 1.9$  arcsec/century) with respect to the BCRS, the largest component of which is the *geodetic (DeSitter-Fokker) precession*, i.e. it would be *dynamically non-rotating* (and Coriolis terms should be added to the equations of motion of bodies in GCRS). Instead in the *kinematically non-rotating* version the *motion of the celestial pole* is defined in GCRS and the *geodetic precession* appears in the *precession-nutation theory* rather than in the transformation between GCRS and BCRS.

In the GCRS of IERS2003 (IERS, 2003) there are small velocities allowing one to use Galilean calculations plus relativistic corrections. However the *ecliptic plane* was redefined only in the IAU2006 resolutions (IAU, 2006) and a nearly relativistic dynamical theory of Earth rotation appeared only in 2009 with IAU2006 (Coppola V., 2009).

Therefore there are still open problems in the relativistic formulation of angular variables (Kaplan G.H., 2005):

- 1) the algorithms for *space motion, parallax, light-time and gravitational deflection* (for the observer at the geo-center the gravity field of the Earth is neglected in evaluating the deflection(star catalogs and ephemerides use 3-vectors in BCRS) );
- 2) the series of rotations for *precession, nutation, Earth rotation and polar motion* (in this order) use 3-vectors in GCRS;
- 3) the *aberration* calculation connects the two systems because it contains the transformation between them: its input are two 3-vectors in BCRS and its output is a 3-vector in GCRS;
- 4) in the VLBI case *aberration* does not appear explicitly, but the conventional algorithm for the *delay* observable incorporates 3-vectors expressed in both systems.

### 3.3 ITRS - International Terrestrial Reference System

The International Terrestrial Reference System ITRS is the Earth-fixed geodetic system which matches the *reference ellipsoid* WGS-84 (basis of the terrestrial coordinates latitude, longitude, height, obtainable from GPS; it has equatorial radius 6,378.137 m and polar flattening  $1/298.257223563$ ) to several centimeters and is defined on the instantaneous 3-spaces  $\Sigma_{t_G}$  of constant geocentric time  $t_G = TCG = const..$  It uses geocentric rectangular 3-coordinates  $\vec{x}_T = \vec{x}_{ITRS}$  on  $\Sigma_{t_G}$  connected to the geocentric ones  $\vec{x}_G$  by time-dependent rotations. It is centered on the geo-center like GCRS with the center of mass defined for the whole Earth including oceans and atmosphere. The coordinates of  $ITRS \approx WGS84 \approx GPS$  are  $(ct_G; \vec{x}_{ITRS})$ . GCRS is obtained from ITRS with a series of *time-dependent rotations* fixed by the conventions in IERS2003 (IERS, 2003) for the precession-nutation theory of Earth rotation. In chapter 5 of Ref. (Kaplan G.H., 2005) there is the old precession-nutation theory, while in chapter 6 there is the new theory of Earth rotation (updated with IAU2006). In Ref. (Coppola V., 2009) there is a more dynamical version IAU2006. Therefore the quasi-inertial relativistic 3-spaces  $\Sigma_{t_G}$  of GCRS are replaced with quasi-Euclidean non-relativistic 3-spaces (still denoted  $\Sigma_{t_G}$ ) only by means of rotations.

The World Geodetic System WGS84 (WGS, 1984) is the latest revision (dated 1984 and revised in 2004) of a standard for use in cartography, geodesy and navigation. It comprises a standard coordinate frame for the Earth, a standard spheroidal reference surface (the reference ellipsoid) for raw altitude data and a gravitational equipotential surface (the *geoid*) that defines

the minimal sea level. The measurement of the form and dimensions of the Earth, the location of objects on its surface and the Earth gravity field are done by means of artificial satellites like the GPS ones (Seeber G., 2003). Let us remark that the gravitational field inside the Earth is evaluated in geodesy with Newtonian gravity, while the external GCRS gravitational potential is evaluated with Post-Newtonian general relativity and the junction of the two approaches has still to be done.

In the description of Earth rotation *precession* and *nutation* are really two aspects of a single phenomenon, the overall response of the spinning oblate, elastic Earth to external gravitational torques from the Moon, Sun and planets. As a result of these torques, the orientation of the Earth's rotation axis is constantly changing with respect to a space-fixed (locally inertial) reference system. The motion of the celestial pole among the stars is conventionally described as consisting of a smooth *long term motion called precession* upon which is superimposed a series of *small periodic components called nutation*.

In the old theory precession and nutation are described by  $3 \times 3$  rotation matrices operating on column 3-vectors in a traditional equatorial celestial coordinate system. The 3-vectors have the

$$\text{form } \vec{x} = \begin{pmatrix} x_x \\ x_y \\ x_z \end{pmatrix} = \begin{pmatrix} d \cos \delta \cos \alpha \\ d \cos \delta \sin \alpha \\ d \sin \delta \end{pmatrix}, \text{ where } \alpha = tg^{-1} \frac{x_y}{x_x} \text{ is the right ascension, } \delta = tg^{-1} \frac{x_z}{\sqrt{x_x^2 + x_y^2}}$$

is the *declination* and  $d$  is the *distance* from the specific origin of the system. For stars and objects *at infinity* (beyond the solar system),  $d$  is often simply put to 1.

In these traditional systems the adjective *mean* is applied to quantities (pole, equator, equinox, coordinates) affected only by precession, while *true* describes quantities affected by both precession and nutation. Thus it is the *true* quantities that are directly relevant to observations; *mean* quantities now usually represent an intermediate step in the computation.

Let us now describe the rotations in the 3-spaces  $\Sigma_{t_G}$  connecting a GCRS 3-vector  $\vec{x}_G$  to a ITRS 3-vector  $\vec{x}_T$  according to the conventions of the new theory of Ref. (IERS, 2003). The new definitions were forced by the fact the the errors in the determination of the old quantities were too big.

A matrix  $\mathbf{B}$ , called *frame bias matrix*, is required to convert ICRS data to the *dynamical mean equator and equinox J2000.0*:  $\vec{x}_{mean(2000)} = \mathbf{B} \vec{x}_{ICRS}$ . The same matrix is used in geocentric transformations to adjust 3-vectors in the GCRS so that they can be operated on by the conventional precession and nutation matrices. The matrix  $\mathbf{B}$  corresponds to a fixed set of very small rotations:  $\mathbf{B} = R_1(-\eta_0) R_2(\xi_0) R_3(d\alpha_0)$  with  $d\alpha_0 = -14.6mas$ ,  $\xi_0 = -16.6170mas$ ,  $\eta_0 = -6.8192$ , all converted to radians (divide by 206264806.247).

If  $\mathbf{B}$  is the frame bias matrix,  $P(t_G)$  the GCRS matrix for precession and  $N(t_G)$  the GCRS matrix for nutation, for a 3-vector  $\vec{x}_G$  in GCRS we have

$$\begin{aligned} \vec{x}_G &\xrightarrow{\mathbf{B}} \text{MEAN EQUATOR and EQUINOX of J2000.0,} \\ &\xrightarrow{P(t_G)} \text{MEAN EQUATOR and EQUINOX at } t_G, \\ &\xrightarrow{N(t_G)} \vec{E}_y = (\text{TRUE EQUATOR and EQUINOX at } t_G). \end{aligned}$$

If the 3-vector  $\vec{E}_Y$  is decomposed on the basis  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ , then by definition the *true equinox at  $t_G$*  in GCRS is the unit 3-vector

$$\vec{Y}_G = B^T P^T(t_G) N^T(t_G) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

By definition the *true celestial pole at date  $t_G$*  - the *Celestial Intermediate Pole CIP* - in GCRS is the unit 3-vector

$$\vec{n}_G = B^T P^T(t_G) N^T(t_G) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \sin d \cos E \\ \sin d \sin E \\ \cos d \end{pmatrix},$$

$$\vec{n}_G \cdot \vec{Y}_G = 0.$$

By definition the *Earth's axis* is the line through the geo-center in direction of the CIP. The angle of rotation about this axis ( $\theta$ , linear in UT1 and *independent from the precession-nutation model for the Earth*) must be measured with respect to some agreed-upon direction in space (CIO, see later on).

The reference point on the equator (origin of  $\theta$ ) must be defined in such a way that the rate of change of the Earth's rotation angle, measured with respect to this point, is the angular velocity of the Earth about the CIP. As the CIP moves, the point must move to remain in the equatorial plane (instantaneously orthogonal to the CIP axis); but the point motion must be such that the measured rotation angle is not contaminated by some component of the motion of the CIP itself. This leads to the concept of *Non-Rotating Origin* (NLO) on the equator: as the equator moves the point's instantaneous motion must always be orthogonal to the equator (whereas the equinox has a motion along the equator: the precession in right ascension). That is, the point motion at some time  $t_G$  must be directly toward or away from the position of the pole of rotation at  $t_G$ . The point is not unique.

The new conventions use the *Celestial Intermediate Reference System CIRS  $E_\sigma$* , which has the NLO azimuthal origin at the *Celestial Intermediate Origin CIO or  $\sigma$* , a well defined point on the equator of CIP with GCRS coordinates

$$\vec{\sigma}_G = \vec{Y}_G \cos \mathcal{E}_o - (\vec{n}_G \times \vec{Y}_G) \sin \mathcal{E}_o, \quad \vec{n}_G \cdot \vec{\sigma}_G = 0,$$

where  $\mathcal{E}_o$  is an angle, named *equation of the origins* (the arc on the instantaneous true equator of date  $t_G$  from the CIO at equinox; it is the right ascension of the true equinox relative to the CIO; it is also the difference  $\theta - \text{GAST}$ , where GAST is the angular equivalent of Greenwich apparent sidereal time), given at p.60 of Ref.(Kaplan G.H., 2005). Now we have an orthonormal triad:  $\vec{n}_G$ ,  $\vec{\sigma}_G$  and  $\vec{Y}_G = \vec{n}_G \times \vec{\sigma}_G$ .

The coordinates in  $E_\sigma$  are

$$\text{GCRS} \xrightarrow{C} E_\sigma, \quad \vec{x}_\sigma = C \vec{x}_G,$$

$$C^T = \begin{pmatrix} \vec{\sigma}_G, \vec{Y}_G, \vec{n}_G \end{pmatrix} = \begin{pmatrix} \sigma_1 & y_1 & x_{CIP} \\ \sigma_2 & y_2 & y_{CIP} \\ \sigma_3 & y_3 & z_{CIP} \end{pmatrix} = R_3(-E) R_2(-d) R_3(E) R_3(s),$$

with the angles  $d$  and  $E$  appearing in  $\vec{n}_G$ . The angle  $s$ , the CIO *locator*, given at p.62 of Ref. (Kaplan G.H., 2005), represents the difference between the length of the arc from the point N westward to the CIO (on the instantaneous equator) and the length of the arc from N westward to the GCRS origin of right ascension (on the GCRS equator).

On the celestial sphere, the Earth's instantaneous (moving) equator intersects the GCRS equator at two nodes. Let N be the ascending node of the instantaneous equator on the GCRS equator.

The matrix  $C$  is the CIO-based rotation taking into account nutation, precession and frame bias.

The Earth rotation angle  $\theta$  with the origin CIO rotates the CIO equatorial axis to an instantaneous axis (the *Terrestrial Intermediate Origin TIO* or  $\tilde{\omega}$ ), which is a NLO azimuthal origin for the *Terrestrial Intermediate Reference System TIRS*  $E_{\tilde{\omega}}$

$$E_{\sigma} \xrightarrow{R_3(\theta)} E_{\tilde{\omega}}, \quad \vec{x}_{\tilde{\omega}} = R_3(\theta) \vec{x}_{\sigma}.$$

Then to arrive to ITRS from  $E_{\tilde{\omega}}$  we must take into account the *polar motion*

$$E_{\tilde{\omega}} \xrightarrow{W^T(t_G)} ITRS, \quad \vec{x}_T = W^T(t_G) \vec{x}_{\tilde{\omega}},$$

$$W(t_G) = R_3(-s') R_2(x_p) R_1(y_p),$$

where  $W^T(t_G)$  is the *polar motion* (wobble) matrix and  $x_p$  and  $y_p$  are the coordinates of CIP in ITRS. This rotation reorients the pole from the ITRS z-axis to the CIP and moves the origin of longitude very slightly from the ITRS x-axis to TIO (the angle  $s' \approx -47 \text{microarcsec}$ ). See p.63 of Ref. (Kaplan G.H., 2005).

In conclusion we have the following sequence of rotations connecting GCRS to ITRS

$$GCRS \rightarrow E_{\sigma} \rightarrow E_{\tilde{\omega}} \rightarrow ITRS,$$

$$\vec{x}_G \rightarrow \vec{x}_{\sigma} = C \vec{x}_G \rightarrow \vec{x}_{\tilde{\omega}} = R_3^T(-\theta) \vec{x}_{\sigma} \rightarrow \vec{x}_T = W^T \vec{x}_{\tilde{\omega}},$$

$$\vec{x}_T = W^T R_3^T(-\theta) C \vec{x}_G.$$

Therefore ITRS is defined by taking the instantaneous 3-spaces  $\Sigma_{t_G}$  of GCRS and by rotating the 3-coordinates in each 3-space to take into account the rotation of the Earth. However in this way all the clocks on the Earth surface have the same geocentric time  $t_G$ . A more relativistic formulation should replace the final rotation matrix  $R$  with a Lorentz transformation  $\Lambda = B R$ , where the Lorentz boost  $B$  would imply the transformation of the global GCRS time  $t_G$  into the different local coordinate times associated with the proper times (SI atomic seconds) of the atomic clocks in each point of the Earth surface.

#### 4. The space-time outside the Solar System

Reference data for positional astronomy, such as the the data in astrometric star catalogs or barycentric planetary ephemerides, are specified in the International Celestial Reference

System ICRS (Kaplan G.H., 2005; Kovalevski J. et al, 1989; Sovers O.J. et al, 1998; Johnstone K.J. et al, 1999; Ma C. et al, 1998; Fey A. et al, 2009) with origin in the solar system barycenter and with spatial axes fixed with respect to space. A materialization as a ICRF is obtained by supposing that the origin is a quasi-inertial observer and that we have a quasi-inertial (essentially non-relativistic) reference frame with rectangular 3-coordinates (or equatorial geographical coordinates) in a nearly Galilei space-time whose 3-spaces are Euclidean. The directions of the spatial axes are effectively defined by the adopted coordinates (i.e. using the tabulated *right ascensions* and *declinations* and, in the case of a star catalogue, the *proper motions* (*ephemerides*)) of 212 extragalactic radio sources observed by VLBI. These radio sources (quasars and AGN, active galactic nuclei) are assumed to have no observable intrinsic angular momentum. At low accuracy one uses a star catalogue system such as the FK5 (Fey A. et al, 2009). At a more accurate level taking into account optical wavelengths, one has the Hipparcos Celestial Reference Frame HCRF, composed of the positions and proper motions of the astrometrically well-behaved stars in the Hipparcos catalog.

Thus, the ICRS is a *space-fixed* system, more precisely a *kinematically non-rotating* system, without an associated epoch. ICRS provides the orientation of BCRS and closely matches the conventional dynamical system defined by the Earth's mean equator and equinox of J2000.0: the alignment difference is at the 0.02 arcsecond level, negligible for many applications.

However if we take into account the description of the universe given by cosmology, the actual cosmological space-time cannot be a nearly Galilei space-time but it must be a cosmological space-time with some theoretical *cosmic time*. In the standard cosmological model (Bartelmann M., 2010; Bean R., 2009) it is a homogeneous and isotropic Friedmann-Robertson-Walker space-time whose instantaneous 3-spaces have nearly vanishing internal 3-curvature, so that may locally be replaced with Euclidean 3-spaces as it is done in galactic dynamics. However they have a time-dependent conformal factor (it is one in Galilei space-time) responsible for the Hubble constant regulating the expansion of the universe. Moreover the Hubble constant is also the negative of the trace of the external 3-curvature of the 3-space as 3-sub-manifold of the space-time. As a consequence the transition from the astronomical ICRS to an astrophysical description taking into account cosmology is far from being understood.

## 5. Concluding remarks

As we have seen relativistic metrology is a field in rapid evolution and subject to continuous refinements.

The existing standard of length will survive till when the technology will allow us to detect the deviations from Euclidean instantaneous 3-spaces implied by Post-Newtonian general relativity. For instance in Ref. (Turyshev S.G. et al, 2006) there is a proposal of a space mission LATOR (Laser Astrometric Test Of Relativity), in which two spacecrafts behind the Sun will form a triangle with the International Space Station ISS. This would allow us to measure the three angles of the triangle to see whether their sum is  $2\pi$  as required by an instantaneous Euclidean 3-space.

The development of optical atomic clocks will allow us to develop a new generation of gravimeters for the local study of the gravitational field of the Earth (now also investigated with the satellites GOCE (Gravity field and steady-state Ocean Circulation Explorer, ESA), CHAMP (CHALLENGING Mini-Satellite Payload, GeoForschungsZentrum GFZ), GRACE (Gravity Recovery and Climate Experiment, Center for Space Research, Austin Texas)). One

open problem to get a reliable theory of heights over the reference geoid is the comparison of the measurements of gravimeters on the two sides of an ocean. But a byproduct of the ACES mission will be the possibility of such a comparison, by synchronizing the optical atomic clocks of the gravimeters with the ACES clocks on the International Space Station ISS (Svelha D. et al, 2008). As a consequence standard non-relativistic geodesy will be replaced by relativistic geodesy.

Also the transformation from the non-relativistic ITRS on the Earth surface to the relativistic GCRS around the Earth will be accomplished. This will put full control on possible semi-relativistic precessional effects near the Earth surface.

Space navigation inside the Solar System will require refinements of BCRS. In particular to test deviations from Einstein theory of general relativity (the one used in BCRS). See for instance the recent interest in the Pioneer anomaly (Turyshev S.G. et al, 2010) and the endless number of proposals for its explanation.

Regarding ICRS we need a general relativistic version of it taking into account the non-Euclidean nature of the 3-space as 3-sub-manifolds of space-time. The unsolved problems of dark energy and dark matter, required by the standard  $\Lambda$ CDM cosmological model starting from the hypothesis of homogeneity and isotropy of space-time, are pushing towards inhomogeneous cosmological space-times in which the 3-spaces have small internal 3-curvature but a non zero external 3-curvature. The first step will be to face these problems inside the Milky Way finding a *relativistic* galactic celestial reference frame extending the existing BCRF. To this end the GAIA (Global Astrometric Interferometer for Astrophysics) mission of ESA (Jordi C., 2011; Jordan S., 2008; Klioner S.A. et al, 2005), to be launched in 2012, for the 3-dimensional cartography of our galaxy (position, proper velocity, radial velocity and spectroscopic data for about one billion stars) will be a first relevant step.

## 6. References

- [1] ACES 2010 - ESA: Atomic Clock Ensemble in Space (ACES), [http : //www.esa.int/SPECIALS/HSF\\_Research/SEMJSK0YDUF\\_0.html](http://www.esa.int/SPECIALS/HSF_Research/SEMJSK0YDUF_0.html).
- [2] Alba D. and Lusanna L., 2010, *Charged Particles and the Electro-Magnetic Field in Non-Inertial Frames: I. Admissible 3+1 Splittings of Minkowski Spacetime and the Non-Inertial Rest Frames*, Int.J.Geom.Methods in Physics 7, 33 (2010) (arXiv 0908.0213) and *II. Applications: Rotating Frames, Sagnac Effect, Faraday Rotation, Wrap-up Effect*, Int.J.Geom.Methods in Physics, 7, 185 (2010) (arXiv 0908.0215).
- [3] Alba D. and Lusanna L., 2007, *Generalized Radar 4-Coordinates and Equal-Time Cauchy Surfaces for Arbitrary Accelerated Observers*, Int.J.Mod.Phys. D16, 1149 (2007) (arXiv gr-qc/0501090).
- [4] Arias E.F., 2005, *The Metrology of Time*, Phil.Trans.R.Soc. A363, 2289 (2005).
- [5] Ashby N., 2003, *Relativity in the Global Positioning System*, Living Rev. Relat. 6, 1 (2003); see also [http : //www.usno.navy.mil/USNO/time/gps](http://www.usno.navy.mil/USNO/time/gps).
- [6] Bartelmann M., 2010, *The Dark Universe*, Rev.Mod.Phys. 82, 331 (2010) (arXiv 0906.5036).
- [7] Basri S.A., 1965, *Operational Foundation of Einstein's General Theory of Relativity*, Rev.Mod.Phys. 37, 288 (1965).
- [8] Bean R., *TASI 2009. Lectures on Cosmic Acceleration* (arXiv 1003.4468).
- [9] Bize S., Laurent P., Abgrall M., Marion H., Maksimovic I., Cacciapuoti L., Grünert J., Vian C., Pereira dos Santos F., Rosenbusch P., Lemonde P., Santarelli G., Wolf P., Clairon

- A., Luiten A., Tobar M. and Salomon C., 2005, *Cold Atom Clocks and Applications*, J.Phys. B38, S449 (2005), special issue *Atoms, Quanta and Relativity: a Century after Einstein's Miraculous Year*.
- [10] Blanchet L., Salomon C., Teyssandier P. and Wolf P., 2000, *Relativistic Theory for Time and Frequency Transfer to Order  $1/c^3$* , Astron.Astrophys. 370, 320 (2000).
- [11] Bonvin C. and Durrer R., 2011, *What Galaxy Surveys Really Measure*, (arXiv 1105.5280).
- [12] Brumberg V.A., 1991, *Essential Relativistic Celestial Mechanics* (Adam Hilger, Bristol, 1991).
- [13] Cacciapuoti L. and Salomon C., 2007, *ACES:Scientific Objectives and Mission Status*, Nucl.Phys. Proc.Suppl. B 166, 303 (2007).
- [14] Cacciapuoti L. and Salomon C., 2008, *Cold Atom Clocks in Microgravity: the ACES Mission*, J.Jpn.Soc. Microgravity Appl. 25, 237 (2008).
- [15] Carroll B.W. and Ostlie D.A., 2007, *An Introduction to Modern Astrophysics* 2nd Ed. (Pearson Education Inc., NewJersey, 2007).
- [16] Chou C.W., Hume D.R., Koelemeij J.C.J., Wineland D.J. and Rosenband T., 2010a, *Frequency Comparison of Two High-Accuracy  $Al^+$  Optical Clocks*, Phys.Rev.Lett. 104, 070802 (2010).
- [17] Chou C.W., Hume D.B., Rosenband T. and Wineland D.J., 2010b, *Optical Clocks and Relativity*, Science 329, 1630 (2010).
- [18] Circular T 263 of the BIPM, 2009, *ftp : //ftp2.bipm.org/pub/tai/publication/cirt.263*.
- [19] Ciufolini I. and Wheeler J.A., 1995, *Gravitation and Inertia* (Princeton Univ.Press, Princeton, 1995).
- [20] Coppola V., Seago J.H. and Vallado D.A., 2009, *The IAU2000A and IAU2006 Precession-Nutation Theories and their Implementation*, AAS 09-159 ([http : //www.centerforspace.com/downloads/files/pubs/AAS-09-159.pdf](http://www.centerforspace.com/downloads/files/pubs/AAS-09-159.pdf)).
- [21] Damour T., Soffel M.H. and Xu C., 1991, *General Relativistic Celestial Mechanics.1. Method and Definition of Reference Systems*, Phys.Rev. D43, 3273 (1991).
- [22] Damour T., Soffel M.H. and Xu C., 1992, *General Relativistic Celestial Mechanics. 2. Translational Equations of Motion*, Phys.Rev. D45, 1017 (1992).
- [23] Damour T., Soffel M.H. and Xu C., 1993, *General Relativistic Celestial Mechanics. 3. Rotational Equations of Motion*, Phys.Rev. D47, 3124 (1993).
- [24] Damour T., Soffel M.H. and Xu C., 1994, *General Relativistic Celestial Mechanics. 4. Theory of Satellite Motion*, Phys.Rev. 49, 618 (1994).
- [25] Durrer R., 2011, *What do we Really Know about Dark Energy?*, (arXiv 1103.5331).
- [26] Fey A., Gordon G. and Jacobs C. (eds.), 2009, *The Second Realization of the International Celestial Reference Frame by Very Long Baseline Interferometry*, IERS Technical Notes 35.
- [27] Finkleman D., Allen S., Seaman J.H. and Seidelmann P.K., 2011, *The Future of Time: UTC and the Leap Second*, American Scientist 99, 312 (2011).
- [28] Garret K. and Duda G., 2011, *Dark Matter: A Primer*, Adv.Astron. 2011, 968283 (2011) (arXiv 1006.2483).
- [29] Gill P., 2005, *Optical Frequency Standards*, Metrologia 42, S125 (2005).
- [30] Guinot B., 1997, *Application of General Relativity to Metrology*, Metrologia 34, 261 (1997).
- [31] IAU 2006 General Assembly (<http://www.iau2006.org/mirror/www.iau.org/iau0603/index.html>).
- [32] IBWM International Bureau of Weights and Measures, 2006, *The International System of Units SI*, p.126 ([http : //www.bipm.org/utills/common/pdf/si\\_brochure\\_8\\_en.pdf](http://www.bipm.org/utills/common/pdf/si_brochure_8_en.pdf))



- [33] *IERS Conventions (2003)*, eds. McCarthy D.D. and Petit G., IERS TN 32 (2004), Verlag des BKG.
- [34] ITUR International Telecommunications Union Recommendation ITU-R TF.460-6, 2007, *Standard-Frequency and Time-Signal Emissions*.
- [35] Johnstone K.J. and de Vegt Chr., 1999, *Reference Frames in Astronomy*, Annu. Rev. Astron. Astrophys. 37, 97 (1999).
- [36] Jordan S., 2008, *The GAIA Project: Technique, Performance and Status*, Astron.Nachr. 329, 875 (2008) (DOI 10.1002/asna.200811065).
- [37] Jordi C., 2011, *The European Space Agency Gaia Mission: Exploring the Galaxy*, (arXiv 1105.6166).
- [38] Kaplan G.H., 2005, *The IAU Resolutions on Astronomical Reference Systems, Time Scales and Earth Rotation Models*, U.S.Naval Observatory circular No. 179 (2005) (arXiv astro-ph/0602086).
- [39] Klioner S.A. and Soffel M.H., 2004, *Refining the Relativistic Model for GAIA: Cosmological Effects in the BCRS*, Proc. of the Symposium *The Three-Dimensional Universe with GAIA*, Paris, pp. 305-309 (ESA SP-576, January 2005) (arXiv astro-ph/0411363).
- [40] Kovalevski J., Mueller I.I. and Kolaczek B., 1989, *Reference Frames in Astronomy and Geophysics* (Kluwer, Dordrecht, 1989).
- [41] Lemonde P., Laurent P., Santarelli G., Abgrall M., Sortais Y., Bize S., Nicolas C., Zhang S., Clairon A., Dimarcq N., Petit P., Mann A., Luiten A., Chang S. and Salomon C., 2001 *Cold Atom Clocks on Earth and Space, in Frequency Measurement and Control, Advanced Techniques and Future Trends*, ed.A.N.Luiten (Springer, Berlin, 2001).
- [42] Lindegren L. and Dravins D., 2003, *The Fundamental Definition of 'Radial Velocity'*, Astron.Astrophys. 401, 1185 (2003) (arXiv astro-ph/0302522).
- [43] Ludlow A.D., Zelevinsky T., Campbell G.K., Blatt S., Boyd M.M., de Miranda M.H.G., Martin M.J., Thomsen J.W., Foreman S.M., Jun Ye, Fortier T.M., Stalnaker J.E., Diddams S.A., Le Coq Y., Barber Z.W., Poli N., Lemke N.D., Beck K.M., and Oates C.W., 2008, *Sr Lattice Clock at 1 CE 10–16 Fractional Uncertainty by Remote Optical Evaluation with a Ca Clock*, Science 319, 1805 (2008).
- [44] Lusanna L., 2011, *Canonical Gravity and Relativistic Metrology: from Clock Synchronization to Dark Matter as a Relativistic Effect*, (arXiv 1108.3224).
- [45] Ma C., Arias E.F., Eubanks T.M., Fey A.L., Gontier A.M., Jacobs C.S., Sovers O.J., Archinal B.A. and Charlot P., 1998, *The International Celestial Reference Frame as Realized by Very Long Baseline Interferometry*, AJ 116, 516 (1998).
- [46] McCarthy D.D. and Seidelmann P.K., 2009, *Time: from Earth Rotation to Atomic Physics* (Wiley, New York, 2009).
- [47] *15th Meeting of the General Conference on Weights and Measures, Resolution 2, 1975* (<http://www.bipm.org/en/CGPM/db/15/2/>).
- [48] *17th Meeting of the General Conference on Weights and Measures, Resolution 1, 1983* (<http://www.bipm.org/en/CGPM/db/17/1/>).
- [49] Moyer T.D., 2003, *Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation* (John Wiley, New York, 2003).
- [50] Parker T.E., 2010, *Long-Term Comparison of Caesium Fountain Primary Frequency Standards*, Metrologia 47, 1 (2010).
- [51] Perlick V., 1987, *Characterization of Standard Clocks by means of Light Rays and Freely Falling Particles*, Gen.Rel.Grav. 19, 1059 (1987).

- [52] Perlick V., 1994, *Characterization of Standard Clocks in General Relativity*, in *Semantic Aspects of Space-Time Theories*, eds. Majer U. and Schmidt H.J. (BI-Wissenschaftsverlag, Mannheim, 1994).
- [53] G.Petit and Wolf P., 2005, *Relativistic Theory for Time Comparisons: a Review*, *Metrologia*, **42**, S138-S144, (2005).
- [54] Resolution 1 of the 13th General Conference on Weights and Measures (CGPM) 1956. See the Official Definition of the International Bureau of Weights and Measures (BIPM) in [http : //www.bimp.org/en/si/si\\_brochure/chapter2/2 - 1/second.html](http://www.bimp.org/en/si/si_brochure/chapter2/2-1/second.html).
- [55] Resolution 10 of the XVIth General Assembly of the IAU, Grenoble 1976 ([http : //www.iau.org/static/resolutions/IAU1976-French.pdf](http://www.iau.org/static/resolutions/IAU1976-French.pdf)).
- [56] Reynaud S., Salomon C. and Wolf P., 2009, *Testing General Relativity with Atomic Clocks*, *Space Sci.Rev.* **148**, 233 (2009).
- [57] Rosenbad T., Hume D.B., Schmidt P.O., Chou C.W., Brusch A., Lorini L., Oskay W.H., Drullinger R.E., Fortier T.M., Stalnaker J.E., Diddams S.A., Swann W.C., Newbury N.R., Itano W.M., Wineland D.J. and Bergquist J.C., 2008, *Frequency Ratio of Al<sup>+</sup> and Hg<sup>+</sup> Single-Ion Optical Clocks; Metrology at the 17th Decimal Place*, *Science* **319**, 1808 (2008).
- [58] Ross M., 2010, *Dark Matter: the Evidence from Astronomy, Astrophysics and Cosmology*, (arXiv 1001.0316).
- [59] Seeber G., 2003, *Satellite Geodesy* (Walter de Gruyter, Berlin NewYork, 2003).
- [60] SIGRAV 2006 - See the talks at the *SIGRAV Graduate School on Experimental Gravitation in Space*(Firenze, September 25-27, 2006) ([http://www.fi.infn.it/GGI-grav-space/egs\\_s.html](http://www.fi.infn.it/GGI-grav-space/egs_s.html)) and at the Workshop *Advances in Precision Tests and Experimental Gravitation in Space* (Firenze, September 28/30, 2006) ([http://www.fi.infn.it/GGI-grav-space/egs\\_w.html](http://www.fi.infn.it/GGI-grav-space/egs_w.html)).
- [61] Soffel M.H., 1989, *Relativity in Astrometry, Celestial Mechanics and Geodesy* (Springer, Berlin, 1989).
- [62] Soffel M.H., Klioner S.A., Petit G., Wolf P., Kopeikin S.M., Bretagnon P., Brumberg V.A., Capitaine N., Damour T., Fukushima T., Guinot B., Huang T., Lindegren L., Ma C., Nordtvedt K., Ries J., Seidelmann P.K., Vokroulicky' D., Will C.M. and Xu Ch., 2003, *The IAU 2000 Resolutions for Astrometry, Celestial Mechanics and Metrology in the Relativistic Framework: Explanatory Supplement* *Astron.J.*, **126**, pp.2687-2706, (2003) (arXiv astro-ph/0303376).
- [63] Sovers O.J. and Fanselow J.L., 1998, *Astrometry and Geodesy with Radio Interferometry: Experiments, Models, Results*, *Rev.Mod.Phys.* **70**, 1393 (1998).
- [64] Standish E.M., 1998, *The Time Scales in the JPL and CfA Ephemerides*, *Astronomy Astrophysics* **336**, 381 (1998).
- [65] Svehla D., Rothacher M., Salomon C., Bayerle G., Wickert J., Helm A., Ziebart M. and Dow J., 2008, *Geodesy Part of the ACES Mission: GALILEO on Board the International Space Station*, talk at the ACES Topical Team Meeting *ACES and GNSS-Based Earth Observation and Navigation*, Institute of Astronomical and Physical Geodesy Technische Universitaet, München, Germany.
- [66] UCLA 2007, *The ABC's of Distances*, ([http : //www.astro.ucla.edu/ wright /distance.htm](http://www.astro.ucla.edu/wright/distance.htm)).
- [67] Turyshev S.G., Shao M. and Nordtvedt K.L., 2006, *Science, technology and mission design for the laser astrometric test of relativity (LATOR)* (arXiv gr-qc/0601035).
- [68] Turyshev S.G. and Toth V.T., 2010, *The Pioneer Anomaly*, *Living Rev.Rel.* **13**, 4 (2010) (arXiv 1001.3686).

- 
- [69] WGS World Geodetic System website of the National Geospatial-Intelligence Agency NGA 1984 (<http://www.nga.mil/ProductsServices/GeodesyGeophysics/WorldGeodeticSystem/Pages/default.aspx>).
- [70] Will C.M., 2011, *Finally, Results from Gravity Probe B*, *Physics* 4, 43 (2011).
- [71] Will C.M., 2006, *The Confrontation between General Relativity and Experiment*, *Living Rev. Relativ.* 9, 3 (2006) (arXiv gr-qc/0510072).
- [72] Will C.M., 1993, *Theory and Experiment in Gravitational Physics*, rev.ed. (Cambridge Univ.Press, Cambridge, 1993).

## **Section 6**

### **Biology and Medicine**



# Development of Metrology for Modern Biology

Sang-Ryoul Park, Jun-Hyuk Choi and Ji-Seon Jeong  
*Center for Bioanalysis,  
Korea Research Institute of Standards and Science,  
Republic of Korea*

## 1. Introduction

### 1.1 Why metrology for modern biology

The full scale DNA information obtained from the Human Genome Project has enabled the revolutionary advances in biotechnologies (BT). Information on human genome is equal to a whole blueprint on human body. Although gigantic, such information can be comprehended and handled with the help of modern information technology. With the information, we can fully understand how our bodies are structured and functioning. Consequently, there will be tremendous advances in human health with innovations in modern medicine driven by BT. It is for sure that mankind will enjoy a healthy and extended life term in a way they never grasped before.

The bright sides of the advanced BT come only with dark sides. It is anticipated that disastrous mishaps could occur as we don't have experiences on such new products of BT. The great outcomes of modern BT can be enjoyed only with meticulous assurance on their safety and efficacy. As those directly touch the quality of life, they could be expensive ones that people have to pay over the financial burdens. Then, the fidelity of such claims will have grave importance to many customers. More important, it can greatly jeopardize lives unless all possible safety concerns have been exhaustively confirmed. Therefore, BT products are subject to extensive testing for quality control (QC) as well as for quality assurance (QA). An unavoidable question here is if the results of such testing are believable. There are numerous formats of bioassays that can be applicable to such testing. Some of those assays might be applied to critical tests for QC and QA without appropriate method validation. Consequently, different results can be produced at different laboratories, which will then create conflicts and disbelief among stake holders. Even worse, some assays that lack analytical reliability can be used to blur the undesirable characteristics of properties of some products.

There have been a number of prominent incidents that insufficient or inadequate testing for QC of biological products greatly risked public health. Among them, "tainted heparin" caused at least 80 deaths. Heparin is used as a blood thinner against formation of blood clotting during surgery or dialysis. The FDA's investigation concluded that heparin problem was due to contaminated raw stock from China with which Boxtor Healthcare produced a prevailing portion of world heparin supply. Tight and adequate testing to detect such fatal contaminants before sales would easily avoid such incidents and save many lives. Similar

tragedy occurred with dairy products as baby meal. As a greedy adulteration of dairy products, melamine was added to milk products to fake high protein content in nitrogen analysis. At least 50 thousand babies were harmed by the melamine in their meal. This incident was also a prominent example of public health risked by insufficient or inadequate testing for quality control of biological products. If QC of dairy products was performed by a method capable of distinguishing real proteins and a nitrogen-rich toxic chemical, such disaster could be easily avoided.

Considering even more complicated products of modern BT, the need for rigorousness of such testing is out of question. There are and will be numerous new products from nucleic acids, proteins, and cells that are highly complex in structures and activities. Consequently, a number of novel analytical approaches are being developed to characterize and quantify such complex biological entities. However, what would be the standards for adequateness of such assessments? Proper answers to the question above should be attained before full-scale industrialization of modern biotechnology and eventually their routine commercial applications. In other areas undergoing earlier industrialization, the support of metrology was critical. The similar or even greater support of metrology is needed for industrialization of biotechnologies. In this area, inaccuracy is more critical, and uncertainty in measurement is much greater. Therefore, metrology for biology should be regarded as a critically important area of modern metrology.

## 1.2 Who work in this area

National metrology institutes (NMI) are on the first line in development of new areas of metrology. The efforts of NMIs are organized in international scale by Bureau International des Poids et Mesures (BIPM, [www.bipm.org](http://www.bipm.org)). In 2000, BIPM arranged the first international meeting for the purpose of development of metrology in biology, which then led to creation of Bioanalysis Working Group (BAWG) as a member group of Consultative Committee on Quantity of Matter (CCQM). Since then, CCQM BAWG has been the most significant international organization to develop metrology for biology. Prominent international organizations related in biology have been working together with CCQM BAWG. World Health Organization-National Institute for Biological Standards and Control (WHO-NIBSC, [www.nibsc.ac.uk](http://www.nibsc.ac.uk)) has been very active in CCQM BAWG activities. In addition, US-Pharmacopeia ([www.usp.org](http://www.usp.org)) and International Federation of Clinical Chemistry (IFCC, [www.ifcc.org](http://www.ifcc.org)) have been communicating closely with CCQM BAWG. A number of international measurement comparisons were made to assess international comparability or to investigate and pinpoint existing major problems in bioanalysis. Among them, quantification of an exogenous gene of genetically modified organisms (GMO) firstly attained interests of participants as it rose as an international concern in both food trade and environmental pollution. Accurate quantification of DNA was the first technical target as it is the most fundamental element in modern biology. As a natural advancement, quantification of proteins has become a central topic. Recently, dare movements toward cells or microorganisms are happening. This is a totally new area where technical advancement is rapid, and measurement issues come urgently. For this reason, proper projection and accordingly planning on future is of high importance. Recently, BIPM has published an investigative report (Marriott et al., 2011) on the current status and future directions of metrology development for biology that is of great value to oversee this particular area.

### 1.3 Approaches and technical issues

Establishment of biometrology is technically demanding. Biological materials are mostly large and complex in their chemical structure. Typical dimensions of important biological entities are illustrated in Fig. 1.1. Even with a highly advanced analytical technique such as mass spectrometry, biological materials are not readily quantifiable with a desired level of accuracy. There are substantial technical gaps between the demands and currently available techniques.

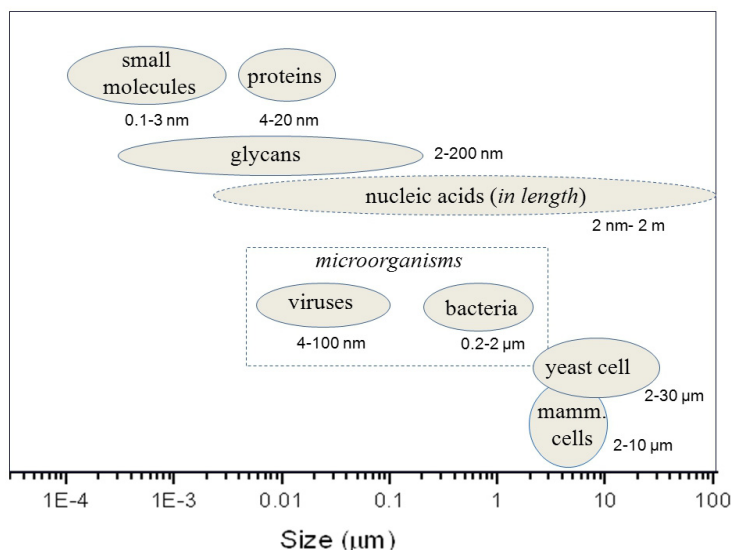


Fig. 1.1. Dimensions (Stokes-Einstein diameters) of biologically important materials. The dimensions of nucleic acids is indicated in length.

Then, approaches to overcome such technical gaps need to be sought out. The approaches suggested by Korea Research Institute of Standards and Science (KRISS, [www.kriss.re.kr](http://www.kriss.re.kr)) scientists as basic approaches for establishment of metrology for biology (Fig. 1.2) would not be much disputable. The first coming idea is to reduce large biological molecules to smaller chemical units that can be accurately quantifiable with current analytical techniques. Reduction to chemical units can be achieved through chemical reactions or enzymatic reactions. The quantity of the original biological molecules is then calculated from the knowledge on their chemical compositions. Uncertainty of such analytical procedures mainly arises from the degrees of perfection of such reduction processes. As far as metrological assessment is made on the uncertainty of the reduction process, the overall analytical procedure can be regarded as of metrological quality. Quantitation of proteins by amino acid composition is a typical example of the reduction approach. Examples of this approach including metrology for protein quantification are described in detail in the following sections.

Large biological materials are discrete and countable. Quantification of cells (cytometry) is an excellent example of count-based quantitation of which history is now longer than



several decades. We can extend this approach whatever biological materials that have a discrete (countable) nature in principle. Quantification of large DNA particles is described in the following section as well as digital-polymerase chain reaction (digital-PCR) that can be also categorized into count-based quantification. In quantification of microorganisms, counting of cultured colonies is of a relatively long history.

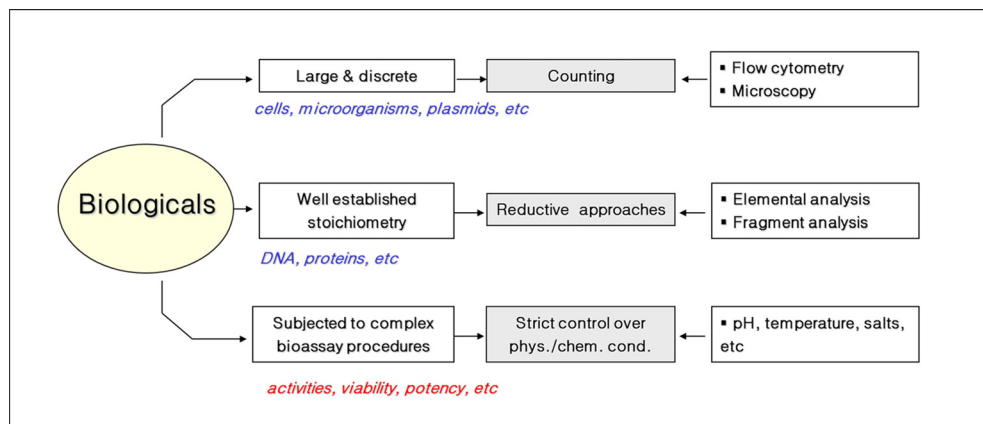


Fig. 1.2. Approaches for establishment of technical basis for metrology for biology suggested by KRISS scientists (from the presentation at 2005 CCQM workshop on “New Challenges for the development of Primary or Higher Order Measurement methods and Procedures”).

Several important biological quantities such as activity, viability, efficacy, and toxicity are not the quantities of biological molecules as above. Instead, these quantities are defined by the experts group of the related fields. These quantities are often called “method-dependent quantities” as those are determined by following the analytical procedures agreed among the experts groups. Metrology for such quantities is therefore of complexity and is often thought to be out of the domain of metrology. However, comparability of the measurement of those quantities is of great importance, and a great deal of improvement can be made by applying metrological rigorously to the physical or chemical conditions involved in each analytical procedure. An example of such aspect is demonstrated in PCR in the following section. Noticeable technical advances have been made in metrology for quantification of nucleic acids and also some for proteins, but not yet for other complex biological materials such as cells and glycans. For these materials, a great deal of research should be done to visualize the practical paths forward realization of metrology. Here, only currently conceivable technical issues are introduced.

## 2. Metrology in quantification of nucleic acids

### 2.1 Backgrounds

Nucleic acids consist of deoxyribose nucleic acids (DNA) and ribose nucleic acids (RNA). It is well known that chemically more stable DNA carries genetic information in long terms whereas RNA promptly response to physiological demands by transferring genetic information to protein productions. For DNA, its base sequences, the genetic codes, are of

primary concerns. Base sequencing of DNA is a full-blown technique these days. The trend in technical advancement was in increment of throughput of analysis. Multiple channel capillary gel electrophoresis has been the most effective and prevailing technique in this regard. Recently, next-generation sequencing (NGS) technology featuring markedly higher throughput than multiple capillary sequencers have been commercialized for more widespread investigation and application of genetic information (Mardis, 2008). Here, the potential bias in DNA sequencing or accuracy in base-calls could be an important issue (Harismendy et al., 2009). However, this area is not in the main scope of metrology as metrology primarily concerns on the quantity. On the other hand, RNA has to be assessed not only about their identities but also regarding their quantities. The expression levels of RNA responding to physiological demands are of critical importance in understanding life processes. Accurate determination of the levels of RNA is especially important to globally share the obtained information. In this regard, related communities such as the Functional Genomics Data Society (FGDS, [www.mged.org](http://www.mged.org)) actively work for achieving satisfactory levels of global harmonization in RNA quantification.

However, accurate assessment of the quantity of DNA is often crucial (O'Connor et al., 2002). One of the simplest forms of DNA is oligonucleotide. Synthetic oligonucleotides are widely used as an essential component of PCR (Dieffenbach et al., 1993; Halford, 1999), probes for various detection schemes (Mitsuhashi et al., 1994), even therapeutic agents as anti-sense drug (Pirolo et al., 2003; Tewary and Iversen, 1997) and agents for RNA interference (Doran, 2004). Some of these applications demand accurate quantitation of oligonucleotides. For example, accurate quantitation of oligonucleotide preparations is one of the most important concerns in the field of DNA-chip technology, where quantitative information from the applied oligonucleotide probes is crucial (Peterson et al., 2001). Accuracy in the quantity of DNA may govern the confidence and inter-comparability of the results of various experiments. In regulations of genetically modified organisms (GMOs), the quantity of a modified gene relative to an endogenous gene is the basis of legal judgments, and technical feasibility in determination of the quantity has been a crucial point for the successful implementation of the legal systems. In modern forensics, revolutionary technical advance has been made by short-tandem repeats (STR)-based human identification. This technique itself is not in the category of DNA quantification, but its reliability is strongly dependent upon the quantity as well as the quality of the obtained DNA samples. Therefore, accurate determination of the quantity of DNA in the sample specimen is an important prerequisite to enable its powerful utility in courts (Brettell et al., 2009). Among various biological substances, nucleic acids come first as the target materials for establishment of metrology for biology. Success in establishment of quantitative metrology for nucleic acids would be the litmus paper for the future success on establishment of metrology for biology.

As discussed in the following section, conventional analytical methods such as UV-absorption and fluorescence measurements as well as recent PCR-based measurements lack measurement traceability to the SI unit of mole. Even the most advanced mass spectrometry (MS) is not suitable to accurately quantify nucleic acids. Numerous reports on applications of MS techniques, mostly matrix assisted laser desorption ionization-time of flight (MALDI-TOF) MS (Bruenner et al., 1996; Zhang and Gross, 2000) are found but are not of adequate accuracy. The complex nature of the ionization of poly-ionic DNA material causes

substantial uncertainty in quantitative analysis. Rigorous evaluations on quantitative performances of mass spectrometry in analysis of polymeric DNA molecules should be further carried out. Meanwhile, other forms of primary analytical methods for establishment of metrology in DNA quantification are demanded. In this regard, notable new approaches are discussed in the following sections.

## 2.2 Conventional methods

Measurement of UV absorbance is most widely used for quantification of DNA. UV absorbance is commonly described in the unit of optical density (OD) of which definition is absorbance of UV light through 1 cm absorption path-length. For DNA, OD for 260 nm ( $OD_{260}$ ) is preferentially used since a local absorption maximum of DNA is at 260 nm. It is known that 1.0  $OD_{260}$  corresponds to the absorption by a single strand oligonucleotide of 30-38  $\mu\text{g/mL}$ , where the most commonly accepted value is 33  $\mu\text{g/mL}$ . Although measurement of UV absorbance is quick and easy, the quantity of DNA estimated from OD is a crude approximation. Bases of DNA consist of four different kinds such as adenine (A), guanine (G), thymine (T) and cytosine (C), and the absorbance of a DNA string varies dependent upon the base composition. The purines of A and G featuring double hetero cyclic rings obviously show stronger molar absorption than the pyrimidine bases of T and C having one ring. The effect of base composition could be reflected in conversion of a UV OD value to a DNA quantity. However, such reflection is not simple due to hypochromicity effects caused by stacking among neighboring bases. A simple summing up method ignoring hypochromicity could lead to as highly as 24% overestimation of the extinction coefficient of a tested oligonucleotide (Cavaluzzi and Borer, 2004). Refinement in consideration of the sequence dependency of the extinction coefficient has been made to 'nearest-neighbor estimates' of extinction coefficients based on mono- and dinucleotide additivity rules (Kallansrud and Ward, 1996). Although quick and convenient, calculation of the concentrations of nucleic acids from the extinction coefficients and UV OD values can't be used as a definitive quantitation method for DNA analysis.

Measurement of fluorescence from dyes intercalated into DNA is known to be a highly sensitive quantitation method of DNA. The dyes are intercalated into DNA proportionally to the number of base units (Yan et al., 1999). Therefore, a linear relationship exists between the intensity of fluorescence and the quantity of the bases. However, the intensity of fluorescence depends on other parameters such as dye concentration, the degree of bleaching, intensity of the excitation light, the optics, and the geometry of a measurement cell as well as the quantity of DNA (Heid et al., 1996). For this reason, calibration using accurate measurement standards is indispensable to obtain accurate quantitative results in fluorometric determination of DNA.

Polymerase chain reaction (PCR), especially in real-time fluorescence detection mode (rt-PCR) or quantitative PCR (q-PCR), is widely used for quantification of specific genes in a DNA mixture for various purposes (Heid et al., 1996; Ponchel et al., 2003). The exponential amplification of PCR renders unprecedented sensitivity to the detection of DNA. If the degree of amplification is monitored at the early stage of the exponential amplification, a linear relationship between the logarithm of the quantity of DNA and the PCR cycle for a given amplification level is obtained (Heid et al., 1996). In PCR amplification, high

selectivity to a specific gene sequence is achieved by using the unique sequences of a primer pair. Therefore, the unknown amount of a target gene sequence can be determined from a calibration curve. However, q-PCR requires calibration standards as the amplification yield substantially depends upon various experimental conditions. Calibration materials of accurately determined quantity are not available yet, and the calibration curves are usually on the scale of the relative quantity of DNA. Although less precise in quantification (~ 10% RSD) and needs to be carefully performed (Halford, 1999), q-PCR is a prevailing quantification method for specific genes and requires urgent development of suitable calibration standards to have measurement traceability to the SI unit of mole.

### 2.3 Reductive approaches

Nucleic acids are polymers of nucleotide units which consist of base, carbohydrate, and phosphodiester bond (Fig. 2.1). As polymeric nucleic acids are not readily quantifiable even with advanced mass spectrometry due to the multiple charges, the components of nucleic acids could be measurement targets in reductive approaches. Firstly, it is noticeable that each unit of nucleotide bears one phosphorus atom in the phosphodiester backbone. Therefore, the quantity of a DNA sequence of a known length (base pairs: bp) can be accurately determined if the quantity of residual phosphorus is quantified accurately. Isotope dilution-mass spectrometry (IDMS) accepted by the metrology community is not an applicable option for determination of phosphorus as the isotope of phosphorus is radioactive ( $^{32}\text{P}$  and  $^{33}\text{P}$ ). Instead, inductively coupled plasma-optical emission spectroscopy (ICP-OES) may be chosen as it is capable of quantification of elemental phosphorus with reasonable precision and accuracy. KRISS scientists first tried this measurement approach where all substantial validation points were carefully examined (Yang et al., 2004). Phosphorus measured by ICP-OES should all come from the target DNA sequence. As a synthetic 20-mer DNA oligomer was quantified as a model DNA, all possible phosphorous contamination in the form of small molecules were removed by applying ultrafiltration of the sample solution. Complete removal was confirmed by the observation of complete disappearance of excessively spiked cytidyl monophosphate (dCMP) after repeated ultrafiltration. In phosphorus measurement, a certified reference material (CRM) of pure phosphorus was used as a calibrant, which provides measurement traceability. With the help of acid and microwave heating, DNA in the sample solution was digested to inorganic phosphates to avoid potential influences from the organic residues of the original molecular structure. An organic compound of a similar structure was excessively spiked into the sample solution to confirm the effect of the spiked organic additive disappears through the digestion process. Analytical performance of ICP-OES in measurement of phosphorus in element solutions has been within 0.2% of accuracy. For the nucleic acids it was estimated within 1% of expanded uncertainty. Sequence impurities of DNA oligomer was minimized by polyacrylamide gel electrophoresis (PAGE) purification, and confirmed using MALDI-TOF MS. As far as such validation points are well confirmed, the ICP-OES method seems to be a fine quantitative method for DNA with measurement traceability and assessable measurement uncertainty. This method has been used to determine the quantity of DNA materials for international comparison studies for DNA quantification. Several publications have been made in an attempt to establish the ICP-OES based phosphorous quantification as a primary analytical method for DNA quantification (English et al., 2006; Holden et al., 2007; Brennan et al., 2009).

One can target different moieties in DNA for the reductive measurement approach. Nucleotide could be the quantification target (O'Connor et al., 2002) whereas nucleoside can be also a measurement target (Donald et al., 2005). Phosphodiesterase produces single nucleotide from a DNA sequence with its activity for hydrolysis of phosphodiester bonds at the 5'-position. If the resultant nucleotides are further subject to phosphatase reactions, then nucleosides are produced. Dependent upon the suitability to available analytical methods, either form can be chosen for quantitation.

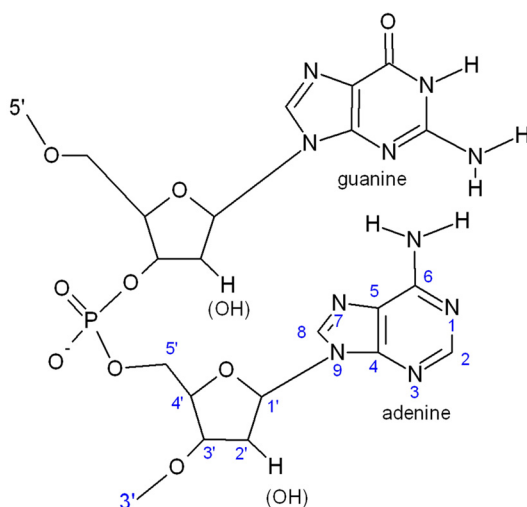


Fig. 2.1. Chemical structure of DNA (or RNA)

If the exact base composition of the target DNA is known, then its quantity can be readily deduced from the results of quantitation of nucleotides or nucleosides. O'Connor et al. first reported the use of HPLC-IDMS for accurate quantitation of nucleotides resulted from enzymatic hydrolysis of a DNA sequence (O'Connor et al., 2002). Later, they reported that analysis of the nucleosides instead of nucleotides resulted in better estimation of the quantity of the target DNA (Donald et al., 2005). For both analytical targets, isotope dilution-mass spectrometry (ID-MS) was applied to secure metrological quality of analysis. However, measurement traceability might not be well established due to the lack of high quality reference materials for preparation of nucleotides or nucleosides standard solutions. Some of those chemicals are too hygroscopic to accurately weigh the actual amounts in preparation of standard solutions. Instead, the exact quantity in the standard solutions can be determined by phosphorous measurement using ICP-OES where any phosphorous contamination of the nucleotide standard solutions can be determined using HPLC-ICP-MS in which different phosphorous-bearing species are separated and determined in terms of phosphorous quantity. Using a high resolution separation method, any partially degraded nucleotides can be also determined. Recently, Hong et al. reported a highly reliable capillary electrophoresis method for determination of nucleotides (Hong et al., 2011). The accurately determined nucleotide standard solutions will provide measurement traceability to this reductive approach. Utilizing those analytical techniques, several NMIs are preparing for the certified reference materials (CRMs) of the nucleotide solutions.

Regardless of the analytical methods applied for analysis of nucleotides or nucleosides, any imperfection of enzymatic hydrolysis will lead to bias in the final results. Therefore, the completeness of the enzymatic hydrolysis has to be carefully assured. Either an insufficient or excessive amount of hydrolysis enzymes may lead to abnormality in the final results. Fortunately, however, any substantial imperfection in hydrolysis can be unambiguously found by comparison of the quantitation results of four nucleotides or nucleosides. This aspect is a great advantage of the methods simultaneously determining four nucleotides or nucleosides. Compared to the ICP-OES method above, this group of analytical procedures requires substantially less amounts of samples, which dramatically expands the applicability. Therefore, the best approach in metrological quantification of nucleic acids will be as follows: 1) measurement traceability is established by the ICP-OES procedure; then 2) quantification of nucleic acids is performed by measuring nucleotides or nucleosides after enzyme hydrolysis where CRMs determined by the ICP-OES procedure is applied.

## 2.4 Count-based quantification

Another interesting approach in metrological quantification of DNA is a count-based quantitation of a trace level DNA. Two methods have been reported in this category. Firstly, several research groups have explored the potential of digital PCR (d-PCR) that aims amplify a single copy of a target gene in a microplate well (Sykes et al., 1992; Vogelstein and Kinzler, 1999). With appropriate dilution, the target gene is to be at the concentration level that only a single copy goes into a microwell or not (Fig. 2.2). Based on Poisson approximation of binomial distribution, at this concentration range, the population of wells that contain a single copy of the target gene among the given number wells will reveal the original concentration of the target gene if the dilution factor is properly considered. Occupation of each well can be determined by the fluorescence signal from the PCR amplicons stained with a fluorescent dye. The concept of digital PCR was first successfully transformed to a commercial instrumentation by Fluidigm ([www.fluidigm.com](http://www.fluidigm.com)), and commercialization was followed by several other companies. Interesting applications of digital PCR can be readily found (Zimmermann et al., 2008).

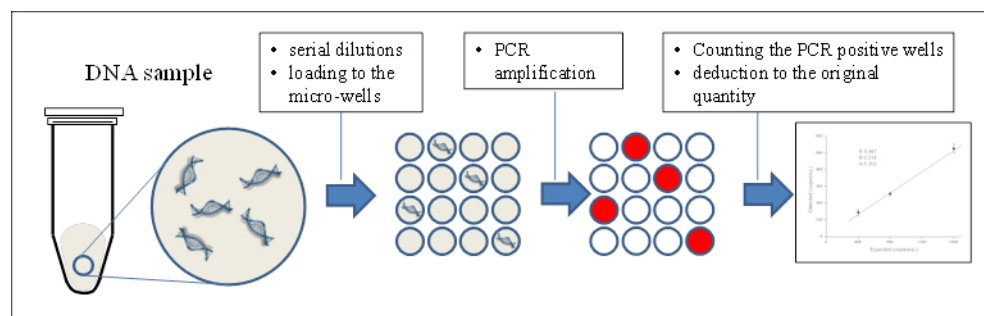


Fig. 2.2. Schematic illustration of the measurement principle of digital PCR for absolute quantification of DNA

Digital PCR, in principle, does not require calibration as the results of quantification comes from the fundamental distribution (Poisson binomial distribution). Therefore, scientists at NMIs pay great attention to this particular measurement technique as it may accurately quantify the amount of DNA without calibration when no appropriate calibration materials are available. A number of publications exploring such possibility can be found (Bhat et al., 2010; Bhat et al., 2009; Corbisier et al., 2010; Sanders et al., 2011). However, there is an important validation point for the use of digital PCR as a metrological DNA quantification method. Calculation for the results of a digital PCR method assumes a 100% success rate for single-copy PCR amplification. Any deviation from 100% success rate will lead to underestimation. In practice, however, 100% success rate for a single-copy PCR is not readily achievable. As reported by Bhat et al., applied PCR conditions such as the status of template DNA or priming sites could result in variations in measurement results (Bhat et al., 2009). For this reason, digital PCR should be carefully validated to draw metrological determination of DNA quantity. Therefore, one would do his or her best to accomplish the 100% success rate for PCR amplification. The judgment for reaching the condition for 100% success rate can be surely made only if a certified reference material (CRM) with actual quantity of the target gene sequence is given.

As described above, digital PCR is a count-based DNA quantitation method for which commercial instruments can be conveniently utilized. However, its validity can be easily checked or demonstrated only if a suitable CRM is available. The other approach of the count-based quantitation of DNA may function in this purpose. KRISS scientists have worked on developing a method and instrumentation for counting individual DNA single molecules on a flow stream. DNA is stained with an intercalating dye and detected by laser induced fluorescence (LIF) detection. For securing enough fluorescence intensity, however, DNA particles need to be large enough. They succeeded in counting lambda phage DNA (45.8 kbp) single molecules (Fig. 2.3) (Lim et al., 2009) and now is capable of counting plasmid DNAs as small as 2 kbp. The LIF detection in this approach does not involve PCR but directly measure fluorescence intensity from the target DNA. Therefore, failure in PCR is not an issue. Direct measurement of fluorescence intensity is rather straight forward and less likely involves errors. In addition, the fluorescence intensity is proportional to the size of the DNA, which gives information on the size of each counted DNA particle. In this counting method, the molar concentration of target DNA is determined by the following simple equation:

$$C = (N_c / N_A) / V$$

Where  $C$ : molar concentration;  $N_c$ : counted number;  $N_A$ : Avogadro's number;  $V$ : sample volume.

As indicated by the equation above, measurement uncertainty may rise from uncertainty in counting of DNA particles ( $N_c$ ) and estimation of sample volume ( $V$ ). There are several check points for assuring accuracy of  $N_c$  as described in detail in the previous publication (Lim et al., 2009). All DNA particles should pass the detection points of a flow channel. Concentration of DNA should be limited under a certain level that simultaneous passing of a number of DNA particles is avoided. Occasional simultaneous passage of two DNA particles should be detectable and counted accordingly by the help of an algorithm for analysis of superimposed peaks. Signal to noise ratio should be large enough to reject background noise in counting. Validation of count-based determination of lambda phage DNA was attempted by comparison with the result of CE-dNMP analysis (Hong et al., 2011), for which the sample concentration was about 100 thousand-fold thicker than the sample for counting. The counting method resulted in about 30% less quantity for lambda phage DNA.

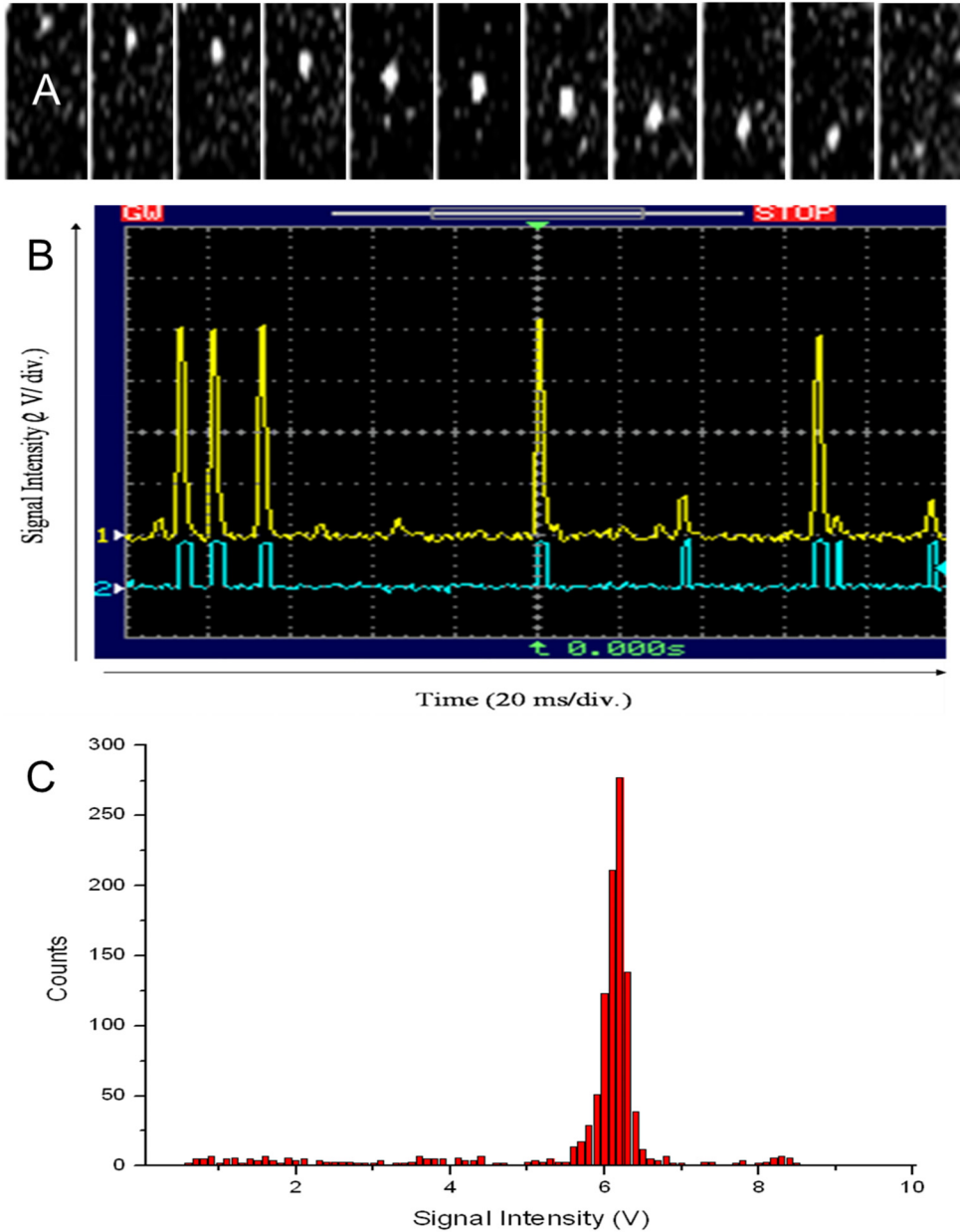


Fig. 2.3. A single lambda DNA captured at 1 ms intervals (A), signals displayed on an oscilloscope (B), and count result displayed in a histogram (C). In panel B, the upper trace indicates the fluorescence intensity, and the lower trace is the digital pulse to be counted (Lim et al., 2009).



In counting plasmid DNA, however, such underestimation disappeared. Signal intensity from plasmid DNA such as pBR322 (4.8 kbp) is 10-fold less than that from lambda phage DNA. Therefore, the underestimation for lambda phage DNA counting should not be due to the weak fluorescence signal for detection. Instead, the underestimation is highly likely due to the fragmentation of lambda viral DNA in the given sample. Small fragments (< 2 kbp) were not counted in counting of lambda phage DNA as the fluorescence signals were far less than that of lambda phage DNA. However, the CE method measures nucleotides from all DNA fragments regardless of their sizes. Therefore, it is possible that the result of CE analysis is substantially greater than the counting result. In contrast, pBR322 DNA that is smaller than lambda phage DNA is too small to be readily fragmented by shear stress (Yoo et al., 2011). The lack of small DNA fragments in pBR322 sample is concordant to the close agreement between the counting method and the CE method. This feature should be carefully considered in determination of what we want to measure. If we want to measure only the DNA molecules of whole integrity, then the counting method gives the right answer from its resultant histogram (count vs. size). On the other hand, if we are interested in the quantity of all DNA fragments, then what the CE method measures is the right answer. Digital PCR in this regard measures a portion of DNA sequences encompassed by a primer pair, which should not necessarily be the same as the number of integral DNA molecules. The “measurand” (exactly what to measure) of each method is significantly different, and this point should be well considered in comparing the results of such methods.

In conclusion, digital PCR will be widely used in attempts to quantify the copies of DNA sequences as commercial instrumentations have become available. However, such applications may have to be confirmed with certified reference materials (CRMs) that are accurately determined by other reliable methods like the direct counting method.

## 2.5 Summary

Metrology for biology has been pursued first in quantification of nucleic acids. Due to the polymeric nature of nucleic acids, conventional methods are not as accurate as a metrological standard should be. In addition, measurement traceability was hardly established. Efforts of many scientists, mostly scientists at various NMIs, have been poured for establishment of metrology for DNA quantification, and substantial achievements have been made. Reductive approaches have been well developed to establish measurement traceability, and count-based quantification methods for trace level DNA copies were also successfully progressed. Although there are some technical challenges still remained, the overall technical structure for metrology in this area is visible now (Fig. 2.4).

Soon, the products of metrology such as certified reference materials in DNA quantification will be available and their widespread uses are expected. Consequently, the credibility in data of DNA quantification will be dramatically improved in various areas such as R&D, industrial, medical, and regulation sectors. More important, the success in the effort to establish metrology for nucleic acid quantification will encourage the scientists who are gazing over the technical challenges in establishment of technical infrastructures for modern biology.

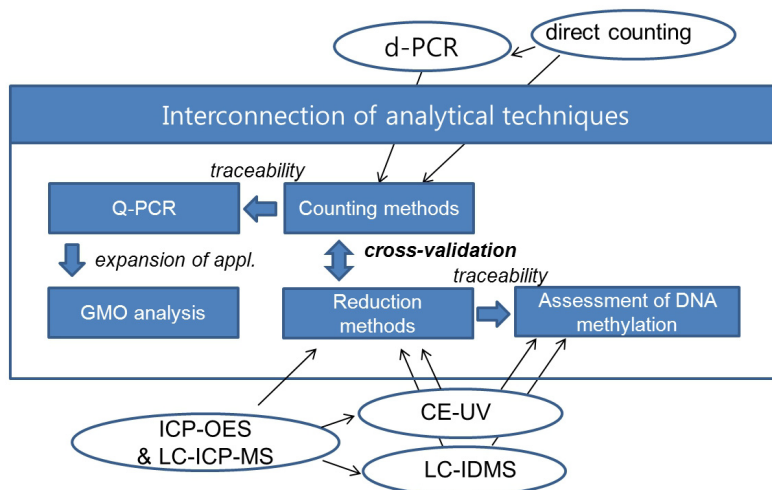


Fig. 2.4. Core analytical techniques developed for establishment of metrology in DNA quantification, and their links to validations or applications

### 3. Metrology in protein quantification

#### 3.1 Backgrounds

Protein is a complex molecule composed of one or more amino acid chains and constitutes a large part of the body's cells, tissues, and organs. This building block of the body has been shown to be a potential biomarker of human health risk assessment. Over the years, therefore, substantial efforts have been made for the development of bioanalytical approaches for accurate quantification of proteins. Monitoring protein levels is important to understand the biological mechanism and regulation of targeted proteins. In clinical practice, such approaches for protein quantification are used to compare biological samples from healthy and diseased patients, providing more accurate diagnosis of human disease. Thus, technologies for accurate quantification of protein biomarkers has been a high priority for many public health related institutions.

Nevertheless, measurement traceability, the essential element of metrology, is hardly established at all in protein quantification. There are several bioassay-based or instrumental methods for protein quantification that are frequently used. All these methods have to be calibrated to produce accurate results, but are not supported with reference calibration materials of metrological grades. Therefore, those methods work at best for comparison of relative amounts. Certified reference materials of proteins to provide measurement traceability to the SI unit of mole is yet to be developed. Our analytical capability is not as advanced as to quantify a whole protein accurately. However, quantification of appropriately reduced forms such as amino acids or peptides using isotopic dilution mass spectrometry (IDMS) should be a feasible approach to quantify proteins in a metrological manner. Of course, such approaches should also be supported by certified reference materials of the reduced forms for accurate calibration. The technical issues in the reduction approaches combined with the use of IDMS are described below.

### 3.2 Conventional methods

There are a number of quantification methods for proteins either bioassays or instrumental methods. The most commonly used methodologies in bioassays are the Biuret (Savory et al., 1968), Bradford (Bradford, 1976), Lowry (Fryer et al., 1986), and bicinchoninic acid (BCA) (Smith et al., 1985) assays. These methods employ chemical reagents which specifically react with proteins to produce colored products which can be measured by UV spectrophotometer in a concentration-dependent manner. The absorbance of colored sample is compared to standard curves constructed with a known protein (frequently bovine serum albumin) in order to determine concentration in unknown samples. Biological methods are widely used owing to their simplicity and low cost, but are hampered by poor accuracy and reproducibility since their responses are either assay or calibrator-protein dependent, and also the results can vary with the residue composition of the target protein. Thus bioassays lead to relative differences in protein quantification and cannot give absolute values. On the other hand, instrumental analytical methods including chromatographic techniques (HPLC, GC), capillary electrophoresis (CE), and mass spectrometry (MS) are often favored because of higher precision. These methods may require time-consuming sample preparation, which leads to higher costs than biological methods. Nevertheless, both bioassay and instrumental analysis need to be calibrated with a highly reliable protein standard material to ensure accurate and comparable results. A higher order analytical method needs to be established for accurate determination of such protein standard materials.

The same is true for high throughput analysis formats such as 2D-PAGE gels (Smithies and Poulik, 1956) and enzyme-linked immunosorbent assay (ELISA). These methods are particularly well fit the purpose of high throughput analysis for proteomics research and protein-chip analysis, respectively. Therefore, they have become powerful tools for screening effective protein markers. As such research progresses, it has become obvious that reliable quantification in an absolute manner is essentially required for comparability of data.

### 3.3 Metrology for protein quantification based on mass spectrometry

Over the last decade, mass spectrometry-based approaches has been shown to be the most powerful tool for protein characterization and become available in many laboratories (Cravatt et al., 2007). Although early biological studies using mass spectrometry merely provided lists of proteins identified in a particular sample, mass spectrometry-based proteomic strategies have been developed, improved, and used for many applications. With the respect to quantitative protein analysis, mass spectrometry alone does not provide much information. MS in general can't detect a whole protein in a quantitative manner. To improve quantification performance, a whole protein needs to be reduced to amino acids or peptides, which are done by chemical or enzymatic hydrolysis, respectively. Even so, repeatability of MS quantification is poor due to the complex nature of ionization prior to MS analysis, which can be dramatically mitigated by the format of isotope dilution-mass spectrometry (IDMS). For the reductive approaches, conversion of the target protein to amino acids or peptides while maintaining the original stoichiometry is a key issue. There could be much debate on the completeness of reduction. This issue is addressed in the following sections. In addition, the sound measurement traceability in quantification of proteins should be established as a very basic requirement for establishment of metrology,

which is highly likely to be achieved through preparation for high quality amino acids CRMs for amino acid based quantification. Nevertheless, the establishment of metrology for protein quantification seems to be a doable task based on advanced mass spectrometry, and the refinements of analytical strategies for reduced measurement uncertainties are pursued as below.

### 3.3.1 Amino acid analysis

Amino acid analysis has been used in many applications as a conventional protein analysis technique. Until the emergence of mass spectrometry, amino acid analysis was essential to the identification of proteins. Nowadays, the analysis is known to be a powerful tool for determination of protein quantities based on detailed information regarding precise quantification of free amino acids. This quantitative analysis is based on the total amount of single amino acids, so the preparation of a highly pure sample is a prerequisite for an accurate analysis. The analysis consists of three steps; hydrolysis of the sample to amino acid constituents (Fig. 3.1), chromatographic separation of the target amino acid to be analyzed, and quantification by mass spectrometry with a labeled internal standard (Fig. 3.2). Of among, complete hydrolysis is critical for overall accuracy of the procedure (Albin et al., 2000; Anderson et al., 1977; Darragh et al., 1996; Fountoulakis and Lahm, 1998; Kinumi et al., 2010). Moreover, for the analysis of an unknown protein, optimization of hydrolysis conditions should be considered before conducting the whole procedure. Very recently, KRIS scientists have reported quantification of human growth hormone by amino acid composition analysis using isotope dilution liquid chromatography-tandem mass spectrometry (Jeong et al., 2011), which is briefly summarized below.

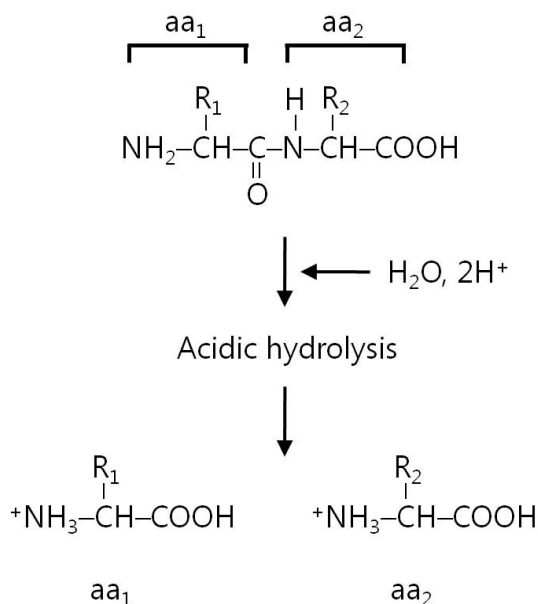


Fig. 3.1. Acidic hydrolysis for amino acid analysis

1. Sample purity assessment: As contaminant components in a protein sample can lead to overestimation of the target quantities, it is important to determine the precise sample purity. To this end, properly diluted human growth hormone (hGH) solution was subjected to capillary zone electrophoresis (CE). The analysis indicates that no significant impurities were present in the sample. The results were also confirmed by high-performance liquid chromatography (HPLC) and matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry.

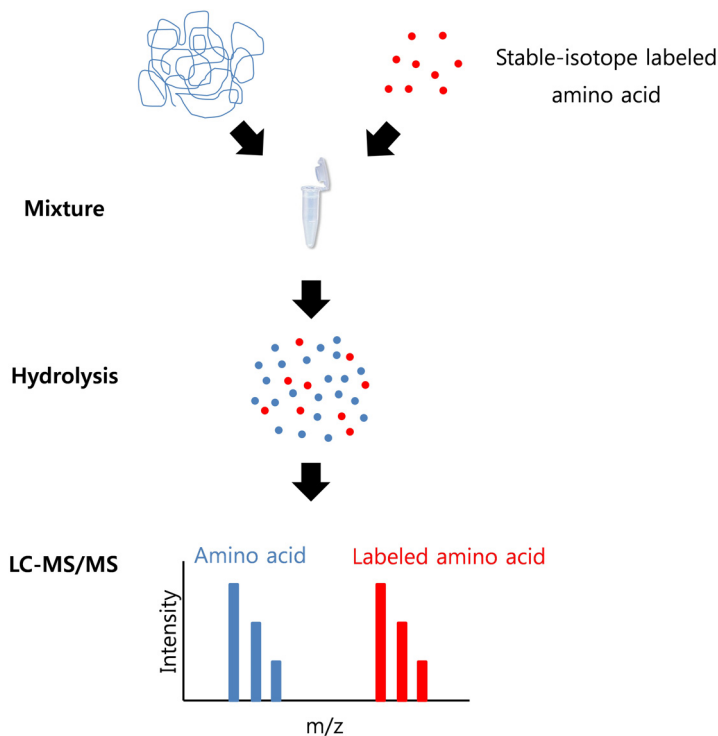


Fig. 3.2. Schematic of MS based amino acid analysis

2. Hydrolysis: As mentioned previously, optimizing the reaction conditions for protein hydrolysis is critical to minimize analytical variations. Through extensive optimization procedures, we found suitable reaction conditions for acidic hydrolysis with hydrochloric acid (HCl), which is the most common hydrolysis method to date. The experiments were performed to test the effects of various factors, including hydrolysis time, protein concentration, HCl concentration, and hydrolysis temperature. As the internal standards should be spiked to the sample prior to the hydrolysis step, the stability of the isotopically labeled free amino acids was also monitored under the identical reaction conditions. The results indicate that the defined hydrolysis step causes no substantial degradation of the internal standard candidates. Therefore, the hGH sample and isotope labeled amino acids were mixed, hydrolyzed, and subjected to HPLC-MS/MS.

3. HPLC-MS/MS: First of all, the optimized conditions for the instrumental analysis were established. Using a high-performance column and a simple isocratic elution, complete separation of the target AAs was observed. It should be noted that full baseline separation of isomeric AAs that might be indistinguishable by MS is often necessary depending on target AAs. In order to perform exact matching double ID-MS quantification, the expected AA concentrations of the hGH sample were calculated, and the same amounts of isotope labeled AAs were used as internal standards. In doing so, issues concerning measurement linearity of MS quantifications could be excluded. MS analysis was performed on a triple-quadrupole mass spectrometer using multiple reaction monitoring (MRM) mode. The ratio of peaks from unlabeled and isotopically labeled AAs was calculated, and doing so the quantity of hGH was determined.

### 3.3.2 Analysis using isotope labeled peptides

In 1980s, a research group reported the preparation and use of a stable isotope-incorporated peptide for measuring endogenous peptides in biological extracts (Desiderio and Kai, 1983; Desiderio et al., 1984). The pioneer work by this group seems to be the first attempt to use isotope labeled synthetic peptides as internal standards to measure the level of a specific peptide using mass spectrometry. After that, another research team described the use of this approach for the absolute protein quantification where a target protein was proteolyzed and quantified using LC-MS/MS with a stable isotope labeled peptide as an internal standard through (Barr et al., 1996). The rapid advances in mass spectrometry based proteomics have greatly improved the analytical sensitivity and made peptide based analysis possible to detect targets in a complex matrix (Anderson et al., 2004; Gerber et al., 2003; Kirkpatrick et al., 2005; Mayya et al., 2006; Putz et al., 2005; Stahl-Zeng et al., 2007). The principal of this approach is that the isotopically labeled synthetic peptide is spiked to the sample, and the mixture is subjected to enzymatic digestion followed by LC-MS/MS. Although this advanced technique is now commonly used for protein quantification, but this approach still has some issues. In addition to high costs of peptide synthesis, significant errors can be generated from sample preparation for mass spectrometry, probably from inefficient and/or inconsistent protease digestion of target proteins. The reaction conditions for protease digestion can be optimized, but the estimated values are likely to be less than the real concentrations of the target proteins. Recently, scientists from multiple NMIs collaboratively investigated the applicability of protein quantification by ID-MS using isotopically labeled synthetic peptides as standards (Arsene et al., 2008), which is briefly summarized below.

1. Purity assessment of the sample and peptide standards: The obtained hGH sample was further purified by liquid chromatography to remove impurities that might interfere with analytical data. Two peptides expected from trypsin digestion of hGH were synthesized in the isotopically labeled forms, where certain AAs were replaced with isotope labeled ones. Importantly, impurities of amino acid and/or peptide in the peptide solutions were carefully evaluated, the resulting factors were corrected.
2. Proteolysis: The defined amounts of the isotope labeled peptides were added to the hGH sample, and then the mixture was subjected to trypsin digestion according to the well-established protocol. In this study, a modified method for fast and complete proteolysis was developed and referred to as the rapid proteolysis protocol. The point of this method is the use of water-organic solvent digestion systems that increase protein solubility and retain high protease activities.

3. LC/MS: After proteolysis by trypsin, the sample containing the internal standards was subjected to LC/MS. The amounts of unlabeled and labeled peptides were monitored by reversed-phase LC/ESI-quadrupole MS using certain ion traces. Comparing the signal ratio of unlabeled and labeled peptides for the standard solution of known concentration to that for the unknown sample, the quantity of hGH was calculated.

### 3.3.3 Analysis using isotope labeled intact proteins

Very recently, isotope-labeled intact proteins have been successfully used for MS based protein quantification (Brun et al., 2007; Janecki et al., 2007; Peng et al., 2004). This approach is based on the use of isotope labeled full-length proteins produced by recombinant DNA techniques as internal standards for quantification. The advantage of this approach is that it is unnecessary to take into account the efficiency of hydrolysis/proteolysis that is the main limitation of amino acid or peptide analysis approaches. As the target analyte and the internal standard are chemically identical, they act exactly the same way during not only pre-analytical treatments such as proteolysis, but also LC/MS measurement. For this reason, this analysis can provide more accurate quantitative analysis than the other MS-based approaches described above, minimizing the influences of experimental variations. However, the difficulty of preparation of isotopically labeled protein has been an obstacle. The following is one representative example of how to quantify a certain protein in a complex biological sample using an isotope-labeled intact protein. Janecki et al accurately measured the level of the human alcohol dehydrogenase (ADH1C1) in liver extracts using in-vitro expressed labeled ADH1C1 (Janecki et al., 2007).

1. Preparation of the stable isotope-labeled ADH1C1: Labeled ADH1C1 was expressed in *Escherichia coli* that were cultured in specific media containing isotopically labeled Lys. The protein was purified by affinity chromatography using His-tags at the N-terminus. The purity and activity of purified ADH1C1 were verified by sodium dodecyl sulfate (SDS)-PAGE and a specific activity assay.
2. Sample preparation: The labeled standard protein was added to liver samples in two different ways. One set referred to as experiment 1 contains various concentrations of the standard in a constant amount of liver extracts. The other set referred to as experiment 2 has a constant amount of the labeled protein in various concentrations of liver extracts. Both experiments were applied to quantification of ADH1C1 in liver extracts to generate statistically significant quantitative data. The mixture was then followed by trypsin digestion, and resulting peptides were recovered.
3. Quantitative analysis by HPLC-MS/MS: Prior to the actual analysis, preliminary experiments were performed using unlabeled ADH1C1 to optimize the multiple reaction monitoring (MRM) transitions, and a specific target peptide was selected. And then, the retrieved peptides from proteolysis were subjected to HPLC-MS/MS using MRM mode. For quantitative purpose, analyst quantification module software was used to calculate the target peaks generated from heavy and normal peptides.

### 3.4 Summary

Determination of the exact amounts of proteins gives valuable information about biological mechanisms of targeted proteins and becomes crucial for the evaluation of various biomarker candidates. Until recently, immunoassays such as ELISA were considered as the

reference methodology for protein quantification. However, these assays have some critical issues such as difficulty of antibody production and non-specific cross-reactivity of antibodies, and thus a significant investment in time and resources is often required for the accurate evaluation. On this regard, the recent development of MS-based protein quantification techniques using isotope labeled internal standards has become an alternative and compelling approach that provides high accuracy and specificity in protein quantification. Although such techniques are now essential for protein quantification, extensive validation of such procedures is necessary in order to establish metrologically sound analytical techniques.

#### 4. Metrology for other biologicals

There are several other important categories of biological materials other than nucleic acids and proteins. They include cells, microorganisms and glycans. One may want to extend the scope to include lipids. Small molecules also play important roles in biological processes as nutrients or metabolites, but it will be better to categorize them as chemicals instead of biologicals. In technical prospects, biologicals here are classified as large molecules that are not readily quantifiable by chemical analysis.

##### 4.1 Cells

A cell is the functional basic unit of life and is much larger and complex than nucleic acids or proteins. Therefore, metrology for cells will be of substantially greater degree of complexity. Recent breathtaking advances in the attempt of utilization of cells in regenerative medicine (Mimeault et al., 2007) draw great attentions of metrologists in biology. Although unconceivable degree of technical difficulties are expected, metrology for cells need to be established to support the advances in regenerative medicine to turn into safe and effective means to restore health of many impaired or elder people. In fact, safety and efficacy of regenerative medicine are debated subjects at present (Ilic and Polak, 2011). In addition, many animal tests are being replaced by cell-based tests to avoid ethical problems, and the validity of cell-based testing will require rigorous examination of cells used for such tests to result in harmonious and valid outcomes. International cell depositories such as American Types of Cell Cultures (ATCC) are the major bodies performing R&D for collective testing of cells, which is necessary for QC and QA of reserved cell resources. In dealing with cells, authentication of the target cells is a key issue, which should be confirmed with effective and valid tests. For example, one of the ATCC R&D programs, Developmental Biology Program (<http://www.atcc.org/Science/CollectionsResearchandDevelopment/DevelopmentalBiologyandStemCells/tabid/207/Default.aspx>), employs state-of-the-art technologies for the authentication of human ES (embryonic stem) cell lines, mouse ES cell lines and mouse feeder cells. Five critical tests address the major concerns of researchers and clinicians working with human ES cell lines and are conducted routinely by ATCC:

1. **Sterility** – confirms the absence of contaminating organisms such as bacteria, fungi, mycoplasma and human viruses
2. **Identity** – confirms species and identity through STR analysis and HLA typing
3. **Stemness** – confirms the cells are in the undifferentiated state using immunocytochemistry, alkaline phosphatase assay, telomerase expression (TRAP), and gene expression (qRT-PCR and global gene expression)



4. **Stability** – confirms the integrity of nuclear genomes using cytogenetics and SNP profile
5. **Pluripotency** – confirms the ability of hESCs to generate various differentiated cell types using teratoma formation and characterization, as well as embryoid bodies formation.

Appropriateness of such tests is to be determined by experts in the field, and the quantities here are more likely method-dependent. As discussed in the later section, metrology for such quantities needs to be established to support ‘metrological execution’ of the given measurement protocols, where those protocols is not fully established yet. In this regard, the advancement of cell testing protocols should be closely tracked for prompt implementation of metrology in such testing protocols.

Counting of well specified cells is rather straight forward in terms of quantification. CCQM BAWG has executed or is preparing for two international comparisons: 1) flow cytometric determination (CCQM-P102) and 2) microscopic determination (CCQM-P123). It is noteworthy that physical parameters such as motility and morphology could be important in characterization of cells. Among NMIs, NIST scientists are in particular active in such investigations (Ni and Chiang, 2007).

## 4.2 Microorganisms

Although there are some useful microorganisms as essentially required in important biological processes such as fermentation, many microorganisms have been life-threatening enemies in long history. Especially, food-born infectious microorganisms are very near to our everyday life to often cause epidemics. The recent *E coli* O157-outbreak in Europe carried by fresh vegetables killed more than 50 people and infected about 4400 people. This incidence reminds that we are not well prepared against such outbreaks. Especially the insufficient technical levels in analysis of microorganisms for surveillance and investigation in a global network is of a great concern. While constructing or improving the global infrastructure against such outbreaks, adopting metrological principles as possible will reduce controversy or legal suits associated with the interpretation on the analytical data. In this regard, CCQM of BIPM moves in a swift manner to organize an international symposium and ad hoc Steering Group on Microbial Measurements to ensure food quality and safety. What to do for establishment metrology in microbial measurements are yet to be rummaged among experts in the field and metrologists. Measurements are for detection, identification, and characterization, and enumeration of target microorganisms. Microorganisms are proliferating dependent upon environmental conditions and frequently mutated. Due to these properties, none of quantification, identification, or characterization of a microorganism is straightforward. Characteristic properties and their relevance on infection should be carefully thought and selected for effective measurement. In addition, such measurements had better be in the domain of metrology, which will guarantee the comparability in a global network (Verhoef et al., 2011). Such a grand task is being organized among relevant experts groups and metrological institutes.

## 4.3 Glycans

Glycan is a polysaccharide, which is a linear or branched polymer of monosaccharide units linked by glycosidic bonds. Glycans exist in nature as free or glycoconjugates (e.g.

glycoproteins, proteoglycans, and glycolipids) with cell or tissue. Their function in living organisms is important. Especially the conjugated glycans in proteins and lipids affect physicochemical properties, such as solubility, stability as well as biological properties. Particularly, glycosylated biopharmaceuticals contain various glycoforms, and the heterogeneity of carbohydrate moiety is altered according to the manufacturing process. For this reason, glycan analysis is an essential part of protein characterization in development and quality control of biopharmaceuticals as well (Harazono et al., 2011). Therefore, the study of protein with post-translational modifications has increased explosively in recent proteomic research, and this phenomenon represents importance of glycans in this area (Wei and Li, 2009). Though the advanced mass spectrometry plays a versatile role in current glycan analysis, due to the complexity of structure and the difficulty of isolation of glycans, there is no single method which can clearly and efficiently analyze all of them (Zaia, 2010). The glycan analysis is processed by a combination of several methods, such as monosaccharide composition analysis after the hydrolysis (Harazono et al., 2011), profiling of liberated glycans (Pabst and Altmann, 2011), and structural assignment (Zaia, 2010). A number of methods have been developed for glycomics using either chromatographic, electromigratory or mass spectrometric method as qualitative and quantitative methods. Glycomics has been recognized as an emerging technology in this post-genome generation in worldwide. The term 'glycomics' includes the systematic study of all glycan structures as well as the comprehensive study in genetic, physiologic and pathologic areas. The dynamic complexity and variety, and the essential roles in physiology, drive the need of generally accepted methodology for glycoconjugates. However, the challenging task to implement metrological rigorousness in such complex analysis is yet to be launched. As new glycomics and glycoproteomics problems will require unique solutions, a wide range of methodological and instrumental techniques will be needed to address them.

## 5. Method-dependent biological quantities

Some important biological quantities are not described in the SI unit of mole. Quantities describing activity, viability, or potency used in biology communities are defined by the methods of assessment that are agreed among the experts in the related fields. Comparability of measurements of these quantities is governed by how precisely the measurement protocols are executed. Therefore, metrology for such quantities needs to be established to support 'metrological execution' of the given measurement protocols. In this regard, improvement of measurement comparability for these quantities can be achieved by somewhat more than providing a simple reference material or reference measurement procedure. In a typical measurement protocol of an enzyme activity, for example, several experimental conditions such as the quantity of the substrate, reaction temperature, pH and buffer composition influence the measurement results. Comparability of measurement of enzyme activity can be assured only when such experimental conditions are tightly controlled. In other words, measurement uncertainties of such quantities arise from the uncertainties in the control of such reaction conditions. It needs to be stressed that such experimental conditions are metrological quantities that can be precisely measured and controlled with clear uncertainty budgeting. Application of metrological rigorousness to such experimental conditions can dramatically improve the measurement comparability for method-dependent quantities. Although yet to be demonstrated, the use of modern computing power will also become essential in dealing with such complex measurement

procedures affected by multiple variables. Highly confusing measurement uncertainties of such procedures may be only correctly budgeted by correct computer- simulations of the procedures.

The potential impact of such notion for method-dependent quantities has been demonstrated in polymerase chain reaction (PCR) (Yang et al., 2005). As described in earlier section, PCR is frequently used as a way to quantify genetic materials, especially in the format of real-time PCR (or q-PCR). In a PCR process, there are several experimental conditions that need to be tightly controlled. Among them, reaction temperature can be a critical control parameter as melting, annealing, and elongation steps of PCR are all profoundly affected by temperature. Numerous commercial thermal cyclers for PCR are available. How well do those thermal cyclers operate if we observe them with a metrological criticism in terms of temperature control? In our investigation (Kim et al., 2008), most PCR instruments showed satisfactory accuracy in static temperature control. However, in dynamic temperature control for PCR, substantially different performances were observed (Fig. 5.1).

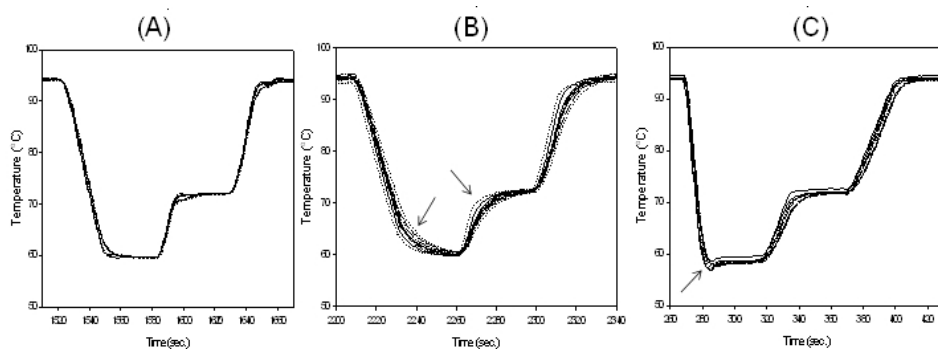


Fig. 5.1. Real-time monitoring of in tube-temperature during PCR cycles for different commercial thermal cyclers. Unlike the ideal case (A), shortened annealing time due to slow response (B), or overshooting (C) and undershooting of temperature were observed (data from Bioanalysis Laboratory at KRISS).

Slower acting instruments did not reach the programmed temperature in a fast PCR. On the other hand, some fast acting instruments showed under- or overshooting, which could lead to drastically different results. Such imperfection of instruments is not well known, and proper measures against such imperfections are not well taken, which would cause substantial disparity in performing bioassays based on PCR. It will be very helpful if the degree of imperfection of instrumentation is analyzed, and is considered in preparation of a measurement protocol. For example, an extreme PCR speed needs to be avoided for users of slower-acting thermal cyclers.

Various bioassays of method-dependent quantity are crucial in QC or QA of biological products. Confidences on biological products are heavily dependent on the credibility and

comparability of such bioassays for QC and QA. Therefore, the establishment of metrology for method-dependent quantities needs to be promptly established. Although brief, an example of how to achieve metrology for such quantities is described above.

## 6. Conclusions

It is certain that metrology for biological measurement will be a critical prerequisite for successful industrialization of modern biotechnology. It will be the fundamental basis of accurate and confident quality control and quality assurance of new products from biotechnology. This notion has been well acknowledged by the international community of metrology, mainly BIPM and CCQM as well as related professional world organizations such as WHO, IFCC, and USP. Because of the extraordinary complexities of biological materials, there were strong doubts about realization of metrology in this area. However, the scientists in national metrology institutes have devoted themselves to construct technical pathways toward the metrology for biology. A number of monumental achievements have been made, and metrology for certain areas especially quantitation of DNA has been established at a highly consistent level. These achievements have a strong symbolic value supporting that we could eventually establish metrology for biology even against many technical difficulties. At this stage, such efforts are now spread toward various important biologicals such as proteins, cells, microorganisms, and glycans. For those biologicals, extensive characterization of exactly what agreed to measure (the measurands) is first coming before quantification, which needs for the attention of not only the experts in metrology but also the experts in the related professional communities. Close communications and active collaborations among scientists in the various disciplines seem to be essential to make significant advancement in metrology for biology. The technical issues in metrology for biology need to be shared among many scientists to invite them to this challenging but interesting area. In this regard, we hope this chapter becomes a little window to look into the world of metrology for biology.

## 7. References

- Albin, D.M., Wubben, J.E., and Gabert, V.M. (2000). Effect of hydrolysis time on the determination of amino acids in samples of soybean products with ion-exchange chromatography or precolumn derivatization with phenyl isothiocyanate. *J Agric Food Chem* 48, 1684-1691.
- Anderson, N.L., Anderson, N.G., Haines, L.R., Hardie, D.B., Olafson, R.W., and Pearson, T.W. (2004). Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J Proteome Res* 3, 235-244.
- Anderson, R., Annette, D., and Jackson, N. (1977). Ion-exchange chromatographic study of amino acid degradation during hydrolysis of avian protein. *J Chromatogr* 135, 447-454.
- Arsene, C.G., Ohlendorf, R., Burkitt, W., Pritchard, C., Henrion, A., O'Connor, G., Bunk, D.M., and Guttler, B. (2008). Protein quantification by isotope dilution mass spectrometry of proteolytic fragments: cleavage rate and accuracy. *Anal Chem* 80, 4154-4160.

- Barr, J.R., Maggio, V.L., Patterson, D.G., Jr., Cooper, G.R., Henderson, L.O., Turner, W.E., Smith, S.J., Hannon, W.H., Needham, L.L., and Sampson, E.J. (1996). Isotope dilution-mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. *Clin Chem* 42, 1676-1682.
- Bhat, S., Curach, N., Mostyn, T., Bains, G.S., Griffiths, K.R., and Emslie, K.R. (2010). Comparison of methods for accurate quantification of DNA mass concentration with traceability to the international system of units. *Anal Chem* 82, 7185-7192.
- Bhat, S., Herrmann, J., Armishaw, P., Corbisier, P., and Emslie, K.R. (2009). Single molecule detection in nanofluidic digital array enables accurate measurement of DNA copy number. *Anal Bioanal Chem* 394, 457-467.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254.
- Brennan, R.G., Rabb, S.A., Holden, M.J., Winchester, M.R., and Turk, G.C. (2009). Potential primary measurement tool for the quantification of DNA. *Anal Chem* 81, 3414-3420.
- Brettell, T.A., Butler, J.M., and Almirall, J.R. (2009). Forensic science. *Anal Chem* 81, 4695-4711.
- Bruenner, B.A., Yip, T.T., and Hutchens, T.W. (1996). Quantitative analysis of oligonucleotides by matrix-assisted laser desorption/ionization mass spectrometry. *Rapid Commun Mass Spectrom* 10, 1797-1801.
- Brun, V., Dupuis, A., Adrait, A., Marcellin, M., Thomas, D., Court, M., Vandenesch, F., and Garin, J. (2007). Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol Cell Proteomics* 6, 2139-2149.
- Cavaluzzi, M.J., and Borer, P.N. (2004). Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res* 32, e13.
- Corbisier, P., Bhat, S., Partis, L., Xie, V.R., and Emslie, K.R. (2010). Absolute quantification of genetically modified MON810 maize (*Zea mays* L.) by digital polymerase chain reaction. *Anal Bioanal Chem* 396, 2143-2150.
- Cravatt, B.F., Simon, G.M., and Yates, J.R., 3rd (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* 450, 991-1000.
- Darragh, A.J., Garrick, D.J., Moughan, P.J., and Hendriks, W.H. (1996). Correction for amino acid loss during acid hydrolysis of a purified protein. *Anal Biochem* 236, 199-207.
- Desiderio, D.M., and Kai, M. (1983). Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. *Biomed Mass Spectrom* 10, 471-479.
- Desiderio, D.M., Kai, M., Tanzer, F.S., Trimble, J., and Wakelyn, C. (1984). Measurement of enkephalin peptides in canine brain regions, teeth, and cerebrospinal fluid with high-performance liquid chromatography and mass spectrometry. *J Chromatogr* 297, 245-260.
- Dieffenbach, C.W., Lowe, T.M., and Dveksler, G.S. (1993). General concepts for PCR primer design. *PCR Methods Appl* 3, S30-37.
- Donald, C.E., Stokes, P., O'Connor, G., and Woolford, A.J. (2005). A comparison of enzymatic digestion for the quantitation of an oligonucleotide by liquid

- chromatography-isotope dilution mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 817, 173-182.
- Doran, G. (2004). Oligonucleotides make sensible strides. *Drug Discov Today* 9, 114-116.
- English, C.A., Merson, S., and Keer, J.T. (2006). Use of elemental analysis to determine comparative performance of established DNA quantification methods. *Anal Chem* 78, 4630-4633.
- Fountoulakis, M., and Lahm, H.W. (1998). Hydrolysis and amino acid composition of proteins. *J Chromatogr A* 826, 109-134.
- Fryer, H.J., Davis, G.E., Manthorpe, M., and Varon, S. (1986). Lowry protein assay using an automatic microtiter plate spectrophotometer. *Anal Biochem* 153, 262-266.
- Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., and Gygi, S.P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc Natl Acad Sci U S A* 100, 6940-6945.
- Halford, W.P. (1999). The essential prerequisites for quantitative RT-PCR. *Nat Biotechnol* 17, 835.
- Harazono, A., Kobayashi, T., Kawasaki, N., Itoh, S., Tada, M., Hashii, N., Ishii, A., Arato, T., Yanagihara, S., Yagi, Y., *et al.* (2011). A comparative study of monosaccharide composition analysis as a carbohydrate test for biopharmaceuticals. *Biologicals* 39, 171-180.
- Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., *et al.* (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 10, R32.
- Heid, C.A., Stevens, J., Livak, K.J., and Williams, P.M. (1996). Real time quantitative PCR. *Genome Res* 6, 986-994.
- Holden, M.J., Rabb, S.A., Tewari, Y.B., and Winchester, M.R. (2007). Traceable phosphorus measurements by ICP-OES and HPLC for the quantitation of DNA. *Anal Chem* 79, 1536-1541
- Hong, N.S., Shi, L.H., Jeong, J.S., Yang, I., Kim, S.K., and Park, S.R. (2011). Rapid and accurate determination of deoxyribonucleoside monophosphates from DNA using micellar electrokinetic chromatography with a cationic surfactant additive. *Anal Bioanal Chem* 400, 2131-2140.
- Ilic, D., and Polak, J.M. (2011). Stem cells in regenerative medicine: introduction. *Br Med Bull* 98, 117-126.
- Janecki, D.J., Bemis, K.G., Tegeler, T.J., Sanghani, P.C., Zhai, L., Hurley, T.D., Bosron, W.F., and Wang, M. (2007). A multiple reaction monitoring method for absolute quantification of the human liver alcohol dehydrogenase ADH1C1 isoenzyme. *Anal Biochem* 369, 18-26.
- Jeong, J.S., Lim, H.M., Kim, S.K., Ku, H.K., Oh, K.H., and Park, S.R. (2011). Quantification of human growth hormone by amino acid composition analysis using isotope dilution liquid-chromatography tandem mass spectrometry. *J Chromatogr A* 1218, 6596-6602.

- Kallansrud, G., and Ward, B. (1996). A comparison of measured and calculated single- and double-stranded oligodeoxynucleotide extinction coefficients. *Anal Biochem* 236, 134-138.
- Kim, Y.H., Yang, I., Bae, Y.S., and Park, S.R. (2008). Performance evaluation of thermal cyclers for PCR in a rapid cycling condition. *Biotechniques* 44, 495-496, 498, 500 passim.
- Kinumi, T., Ichikawa, R., Arimoto, H., and Takatsu, A. (2010). Traceable amino acid analyses of proteins and peptides by isotope-dilution mass spectrometry using precolumn derivatization reagent. *Anal Sci* 26, 1007-1010.
- Kirkpatrick, D.S., Gerber, S.A., and Gygi, S.P. (2005). The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. *Methods* 35, 265-273.
- Lim, H.-M., Yoo, H.-B., Hong, N.-S., Yang, I., Han, M.-S., and Park, S.-R. (2009). Count-based quantitation of trace level macro-DNA molecules. *Metrologia* 46, 375-387.
- Mardis, E.R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet* 24, 133-141.
- Marriott, J., O'Conner, G., Parkes, H. (2011). Final Report: Study on Measurement Services and Comparison Needs for an International Measurement Infrastructure for the Biosciences and Biotechnology: Input for the BIPM Work Programme, Rapport BIPM-2011/02, Reprot Number: LGC/R2011/123 Number, BIPM.
- Mayya, V., Rezual, K., Wu, L., Fong, M.B., and Han, D.K. (2006). Absolute quantification of multisite phosphorylation by selective reaction monitoring mass spectrometry: determination of inhibitory phosphorylation status of cyclin-dependent kinases. *Mol Cell Proteomics* 5, 1146-1157.
- Mimeault, M., Hauke, R., and Batra, S.K. (2007). Stem cells: a revolution in therapeutics--recent advances in stem cell biology and their therapeutic applications in regenerative medicine and cancer therapies. *Clin Pharmacol Ther* 82, 252-264.
- Mitsubishi, M., Cooper, A., Ogura, M., Shinagawa, T., Yano, K., and Hosokawa, T. (1994). Oligonucleotide probe design--a new approach. *Nature* 367, 759-761.
- NI, Y., Chiang, Y.M. (2007). Cell Morphology and Migration Linked to Substrate Rigidity. *Soft Matter* 3, 1285-1292.
- O'Connor, G., Dawson, C., Woolford, A., Webb, K.S., and Catterick, T. (2002). Quantitation of oligonucleotides by phosphodiesterase digestion followed by isotope dilution mass spectrometry: proof of concept. *Anal Chem* 74, 3670-3676.
- Pabst, M., and Altmann, F. (2011). Glycan analysis by modern instrumental methods. *Proteomics* 11, 631-643.
- Peng, J., Kim, M.J., Cheng, D., Duong, D.M., Gygi, S.P., and Sheng, M. (2004). Semiquantitative proteomic analysis of rat forebrain postsynaptic density fractions by mass spectrometry. *J Biol Chem* 279, 21003-21011.
- Peterson, A.W., Heaton, R.J., and Georgiadis, R.M. (2001). The effect of surface probe density on DNA hybridization. *Nucleic Acids Res* 29, 5163-5168.
- Pirollo, K.F., Rait, A., Sleer, L.S., and Chang, E.H. (2003). Antisense therapeutics: from theory to clinical practice. *Pharmacol Ther* 99, 55-77.

- Ponchel, F., Toomes, C., Bransfield, K., Leong, F.T., Douglas, S.H., Field, S.L., Bell, S.M., Combaret, V., Puisieux, A., Mighell, A.J., *et al.* (2003). Real-time PCR based on SYBR-Green I fluorescence: an alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. *BMC Biotechnol* 3, 18.
- Putz, S., Reinders, J., Reinders, Y., and Sickmann, A. (2005). Mass spectrometry-based peptide quantification: applications and limitations. *Expert Rev Proteomics* 2, 381-392.
- Sanders, R., Huggett, J.F., Bushell, C.A., Cowen, S., Scott, D.J., and Foy, C.A. (2011). Evaluation of digital PCR for absolute DNA quantification. *Anal Chem* 83, 6474-6484.
- Savory, J., Pu, P.H., and Sunderman, F.W., Jr. (1968). A biuret method for determination of protein in normal urine. *Clin Chem* 14, 1160-1171.
- Smith, P.K., Krohn, R.L., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* 150, 76-85.
- Smithies, O., and Poulik, M.D. (1956). Two-dimensional electrophoresis of serum proteins. *Nature* 177, 1033.
- Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Krek, W., Aebersold, R., and Domon, B. (2007). High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol Cell Proteomics* 6, 1809-1817.
- Sykes, P.J., Neoh, S.H., Brisco, M.J., Hughes, E., Condon, J., and Morley, A.A. (1992). Quantitation of targets for PCR by use of limiting dilution. *Biotechniques* 13, 444-449.
- Tewary, H.K., and Iversen, P.L. (1997). Qualitative and quantitative measurements of oligonucleotides in gene therapy: Part II in vivo models. *J Pharm Biomed Anal* 15, 1127-1135.
- Verhoef, L., Kouyos, R.D., Vennema, H., Kroneman, A., Siebenga, J., van Pelt, W., Koopmans, M. (2011). An integrated approach to identifying international foodborne norovirus outbreaks. *Emerging Infectious Diseases* 17, 412-418.
- Vogelstein, B., and Kinzler, K.W. (1999). Digital PCR. *Proc Natl Acad Sci U S A* 96, 9236-9241.
- Wei, X., and Li, L. (2009). Comparative glycoproteomics: approaches and applications. *Brief Funct Genomic Proteomic* 8, 104-113.
- Yan, X., Grace, W.K., Yoshida, T.M., Habbersett, R.C., Velappan, N., Jett, J.H., Keller, R.A., and Marrone, B.L. (1999). Characteristics of different nucleic acid staining dyes for DNA fragment sizing by flow cytometry. *Anal Chem* 71, 5470-5480.
- Yang, I., Han, M.S., Yim, Y.H., Hwang, E., and Park, S.R. (2004). A strategy for establishing accurate quantitation standards of oligonucleotides: quantitation of phosphorus of DNA phosphodiester bonds using inductively coupled plasma-optical emission spectroscopy. *Anal Biochem* 335, 150-161.
- Yang, I., Kim, Y.H., Byun, J.Y., and Park, S.R. (2005). Use of multiplex polymerase chain reactions to indicate the accuracy of the annealing temperature of thermal cycling. *Anal Biochem* 338, 192-200.



- Yoo, H.-B., Lim, H.-M., Yang, I., Kim, S.-K., and Park, S.-R. (2011). Flow cytometric investigation on degradation of macro-DNA by common laboratory manipulations. *J. Biophysical Chemistry* 2, 102-110.
- Zaia, J. (2010). Mass spectrometry and glycomics. *OMICS* 14, 401-418.
- Zhang, L.K., and Gross, M.L. (2000). Matrix-assisted laser desorption/ionization mass spectrometry methods for oligodeoxynucleotides: improvements in matrix, detection limits, quantification, and sequencing. *J Am Soc Mass Spectrom* 11, 854-865.
- Zimmermann, B.G., Grill, S., Holzgreve, W., Zhong, X.Y., Jackson, L.G., and Hahn, S. (2008). Digital PCR: a powerful new tool for noninvasive prenatal diagnosis? *Prenat Diagn* 28, 1087-1093.

# Dimensional Measurement for Dentistry

Reisha Rafeek<sup>1</sup>, Kevin Seymour<sup>2</sup> and Lifong Zou<sup>2</sup>

<sup>1</sup>*School of Dentistry, University of the West Indies,*

<sup>2</sup>*Barts and The London School of Medicine and Dentistry,*

*Queen Mary University of London,*

<sup>1</sup>*Trinidad and Tobago*

<sup>2</sup>*England*

## 1. Introduction

Metrology is the science of measurement. Although it has been extensively used in the engineering and manufacturing fields for decades, measurement applications in dentistry are a relatively recent development. One of the major differences of measuring objects in dentistry is that the shape of human body parts are “free-form” rather than designed geometrical shapes used in industry.

## 2. Main issue

Dentistry requires development in hand skills throughout the undergraduate and postgraduate programs both in the pre-clinical and clinical settings (Allred, 1977). The level of expertise achieved depends on the level of training, the natural ability of the dentist and also experience. Tooth preparation for a crown is a common procedure in general dental practice and it is essential that dental students are able to perform this procedure competently before they graduate. The skill training of these techniques is in itself subjective, as there is use of “eyeballing” of the preparation rather than a definite measure. There is clearly a need to obtain more formal dimensional assessment of crown preparations. This will assist in the feedback to and the training of dentists and also in the practise of dentistry.

This chapter will describe dimensional measurements in dentistry with traditional solutions and present the use of metrology in dentistry. Metrology can benefit the teaching of dentistry by presenting three-dimensional images of teeth before and after preparation and measuring dimensional changes. This is also a benefit to the student as they are able to receive more objective measurable feedback. Metrology has also been used in dentistry to measure the amount of tooth wear on the surface of a tooth. Specifically in the preparation of crowns, it has been used previously to measure the width of the crown preparation margin, the angulation of the margin (Seymour et al.,1996, Cherukara et al., 2002) and the taper angle formed when a tooth is prepared for a crown (Rafeek et al., 2006, 2010).

### 2.1 Traditional solutions

The retention of a crown on a prepared tooth not only depends on the length and diameter of the walls of the preparation but also what is called the taper of the preparation. The angle

formed between opposing walls of the tooth preparation for a crown is called taper or convergence angle (Rosenstiel, 1975). The retention of crowns decreases with increasing taper and has been shown to be inversely proportional to taper or convergence angle (Jorgenson, 1955). Therefore the higher the taper, the less retentive the crown and the more likely it will dislodge and come off the patient's tooth which is an undesirable outcome. The taper produced is in the hands of the dentist while the length and diameter of the walls are a function of the existing tooth. The ideal taper recommended is  $4^{\circ}$  -  $14^{\circ}$  total convergence angle (Shillingburg et al., 1997) and some dental schools also recommend similar angles (Noonan & Goldfogel, 1991; Robinson & Lee, 2001). Taper has been measured previously by various methodologies using replicas of the tooth preparations called dies and a microscope (Ohm & Silness, 1978), photocopying (Noonan & Goldfogel, 1991), use of shadowgraphs (Weed et al., 1984; Sato et al., 1998), overhead projectors (Nordlander et al., 1988; Patel et al., 2005) and a goniometer microscope (Ayad et al., 2005).

The replicas of the teeth or dies are made from dental stone and these dies represent the tooth that has been prepared in the patient's mouth. The three dimensional shape has four sides and one top or occlusal surface. Two walls that oppose each other are called mesial and distal and the other two are called buccal and lingual. It is the angle formed by two opposing walls, for example between the buccal and lingual walls (in the bucco-lingual plane) which is measured as Figure 1.



Fig. 1. Showing buccal and lingual walls (left) and mesial and distal walls (right)

Ohm and Silness (1978) projected the external contour lines of the preparation (of nearly 200 dies) and enlarged them onto the screen of the microscope and then measured the convergence angle. Noonan and Goldfogel (1991) used photocopy machines to copy over 900 dies. The dies were placed so that the shadow picture passed through the mesio-distal plane and then the images were produced on the copy paper. Lines were drawn on the surface planes of the shadow picture and the angle of convergence calculated.

Sato et al., (1998) investigated over 60 dies by projecting the silhouettes of the dies in a mesio-distal and bucco-lingual plane with a profile projector at magnification and traced them onto paper. Lines were drawn parallel to the walls and extended until they met and formed the convergence angle which was measured with a protractor. Patel et al., (2005) conducted investigations of convergence angle by also projecting silhouettes of 240

sectioned dies onto a flat surface by means of an overhead projector. This methodology was used previously by Nordlander et al., (1988). The devices such as photocopiers and overhead projectors were inexpensive to use and easily available. However the projecting of the dies to produce silhouettes would produce some error due to a 3-D object being measured as a 2-D image. Also, with the overhead projector there may have been inaccuracy due to the magnification of the image.

Ayad et al., (2005) used a goniometer microscope to measure the convergence angle of silhouettes of almost 500 die specimens in the mesio-distal and bucco-lingual plane. The angle formed with one prepared wall e.g. the buccal wall, and the vertical crosshair within the focusing screen was measured in degrees with the vernier scale on the microscope. The taper or convergence angle was calculated by adding the buccal and opposing lingual measurements. The goniometer microscope is a more expensive device than the ones used previously but more highly accurate.

Several studies conducted using these traditional methodologies investigating tapers achieved by dental students have found mean tapers ranging from 11° - 27° (Noonan & Goldfogel, 1991; Robinson & Lee, 2001; Smith et al., 1999; Mack, 1980; Ohm & Silness, 1978; Sato et al., 1998; Patel et al., 2005; Ayad et al., 2005) while studies on specialists and general dental practitioners and have found mean tapers in the range 14° - 20° (Patel et al., 2005; Nordlander et al., 1988; Kent et al., 1988; Eames et al., 1978). Even though higher than ideal tapers are found in some studies, tapers of up to 20° have been shown to be clinically acceptable with only a few crowns that were reported to have loosened and lost retention or dislodged (Ohm & Silness, 1978; Sato et al., 1998).

## 2.2 Proposed alternative

This alternative to the traditional types of devices for measurement is a modern technology used in manufacturing industry called 3-D co-ordinate metrology. It has been very difficult to accurately measure the taper produced and these previous methodologies may have had errors. The use of the Co-ordinate Measuring Machine (CMM) is an attempt to more formally, consistently and accurately measure dimensional changes due to tooth preparation either at the margin of the crown, the angulation of that margin or the taper of the walls of the preparation. It characterises and defines the geometry of objects.

Free form surface measurement involves two key procedures, data acquisition and development of the correct fitting algorithm (Summerhays *et al.*, 2002). Each of these problems has numerous solutions which largely depend upon the specified application. For example, the application of non-contact measurement methods has expanded into many new areas from traditional industrial applications; to the medical field, such as measurements of the human face and other parts of the human body (Aung *et al.*, 1995; Shahrom *et al.*, 1996; Fan, 1997). Gradually, researchers have required the process to be used on more complicated surfaces such as in the quantification of tooth morphology.

There are number of techniques currently used to obtain accurate surface measurements, each has its own strengths, weaknesses and areas of applicability. With surface form (shape) measurements, Co-ordinate Measuring Machines (CMMs) are more suited for the inspection of automotive parts in industry since the 1950's (Bosch, 1995). Since the late 90's, along with the development of new technology in optics and computing, numerous applications have

been advanced in the fields of medicine and dentistry (Chadwick, 1989; Seymour *et al.* 1996; Kabban, *et al.* 2001; Zou, *et al.* 2001, 2009; Cherukara GP *et al.* 2002; Theocharopoulos *et al.*, 2010; Morris *et al.* 2011 ).

The interface between the CMM and the measurement object or surface is a probe. The probe plays a very important role in obtaining the co-ordinates of a surface, and there are several kinds of probes commercially available at the present time. **Touch trigger probes** are by far the most commonly used with CMMs due to their simplicity and low cost. One of the problems with touch probes is that even under well-controlled conditions and at a generally accepted measuring speed and force, damage to the surface being scanned can occur because of the inertia effects of some probe parts (Win *et al.*, 1998). Win also indicated that initial plastic yielding of a surface can easily occur when probing with a sphere, and at higher probing speeds, multiple bouncing of the probe may occur, which reduces the accuracy for all mechanical probe systems. **The optical probe** has the great advantage of non-contact with the surface being measured. By the late 1980's, non-contact laser probes were widely accepted for surface measurement, especially in free form surface digitisation techniques due to its fast speed in the digitisation. However the accuracy in measurement when using optical probes is affected by the reflectivity of the surface, for example, a rough or dark coloured surface which is often encountered, results in empty or distorted data readings at areas due to the surface low reflectivity.

### 2.2.1 Surface digitisation technology

Surface digitisation is a procedure by which digital data is acquired from a surface of interest in order to obtain more knowledge of the shape and dimensions of the object. This is achieved through a series of interactions between the measurement equipment and the surface being assessed. Digitisation usually provides a string of 3-D co-ordinates in ASCII format.

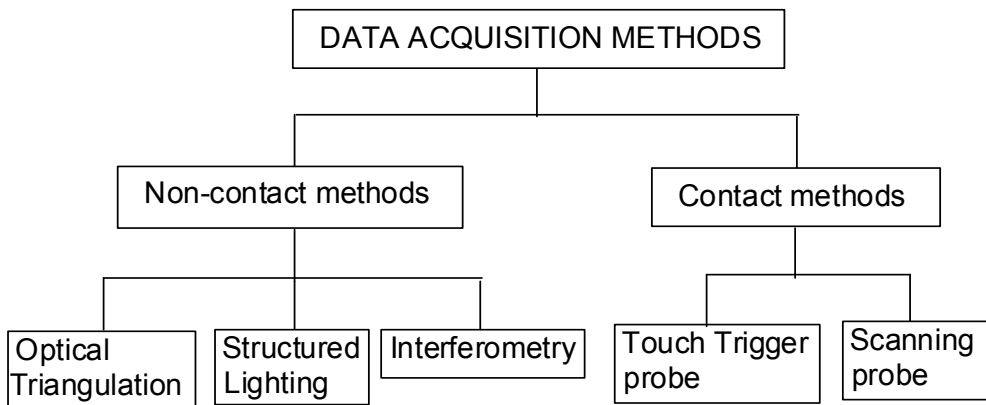


Fig. 2. A classification of data acquisition methods.

There are many different methods for acquiring co-ordinate data, as shown in Figure 2. Essentially, each method uses some mechanism or phenomenon for interacting with the surface of the object of interest.

Contact methods use mechanical contact through a trigger or scanning probe at the end of an arm or the column of a CMM (Jarvis, 1983) to obtain surface geometrical information. Development of this method was driven by reverse engineering (Milroy *et al.*, 1996; Dalton, 1998; Chow, 1997). In each case, an appropriate sensor determines a precise position on the object's surface.

Non-contact methods use scanning-type CMMs to capture larger numbers of sampling points, and have been used successfully for measuring surface form (Bradley and Vickers, 1992; Lee *et al.*, 2001; Chang and Lin, 1999). A recent study comprising questionnaire respondents from research institutions and companies engaged in surface topography in the European Community, revealed that non-contact optical techniques are more popular and widely used than previously thought (Dong *et al.*, 1994). Three dimensional non-contact measurement methods started to be used in the 1980's (Bosch, 1995; Stout, 1994).

The result of data acquisition is a discrete 'point cloud', through a 3-D analysis software it is interpolated into a 3-D surface and can be visualised as a 3-D object, furthermore to be analysed in its shape and dimensions.

With regard to the previously mentioned taper or convergence angle of prepared teeth; an angle measurement is taken at the middle of the mesial-distal plane where the central fissure runs through. The taper is the angle between the two lines along the two walls in this plane that passed the most outer points A and B which are located at the mesial and distal aspects as shown in Figure 3.

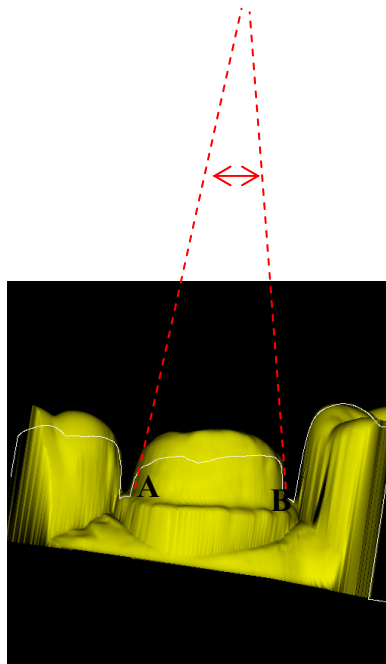


Fig. 3. The angle measurement between mesial and distal walls

### 2.3 Experimental measurements

Experimental studies were conducted in order to assess the taper of teeth prepared for crowns by dental students. Teeth were prepared both in the clinical setting on live patients and also on artificial plastic teeth set in a manikin phantom head jaw in a preclinical laboratory. It would then not only be possible to measure taper achieved but to also compare this with other dental schools.

Dental students are first taught to carry out crown preparations in a laboratory based course with didactic sessions of lectures and handouts. There are also practical sessions with live demonstrations projecting magnified images on television monitors via video cameras of the procedure in detail. They are taught to create a preparation with a 6° taper by holding a 3° tapered bur parallel to the long axis of the tooth while cutting each wall (Shillingburg et al., 1997). The students are then expected to perform the procedures until satisfactory and then pass practical examinations for that procedure prior to being allowed to perform the procedure on a patient.

The specimens used were obtained from both crown preparations carried out by students for their practical examinations in the laboratory and also from crown preparations done in the clinic on patients. There were 99 laboratory specimens or dies and 40 clinical dies. In order to form the replicas or dies of the tooth preparations an impression of the tooth has to be taken. This was done by using addition cured polyvinylsiloxane impression material, Repronil (Dentsply Caulk, Milford DE, USA). The technician who fabricated the crowns used the first pour of the impression to fabricate the actual crown that was fitted onto the tooth. The impression can be used again for a second pour in order to make the die that was used to measure the taper. This second use of this type of impression has been shown to be as accurate as the first pour (Johnson & Craig, 1985). The die stone used to pour into the impression to create the die replica was Type IV (Vel Mix) (Kerr Corp., Orange CA, USA) and was mixed under vacuum and poured. The material was left to set and the standing base of the dies were trimmed specifically to be flat to the long axis of the prepared tooth and the sides of the base were also milled so that the buccal and lingual were parallel to each other and the mesial and distal parallel to each other.

The die spacer (Quick Set Die Spacer Blue, Belle de St. Claire, (Kerr Corp., Orange CA, USA) was painted over the crown preparations in 2 layers according to manufacturer's instructions, as is the usual procedure when fabricating crowns. This also overcame the incompatibility of the pink stone to the probe and was needed in order for the scanning machine to effectively read the preparation, allowing the probing tracer to follow the surface profile rather than the details of the surface texture.

The die stone replica was fixed in a clamping device and digitised by the purpose built dental scanner Incise (Renishaw, UK) with a scanning probe of SM25-1 (Renishaw, UK) shown in Figure 4. The digitisation is conducted by Tracecut Controlling Software Package (Renishaw, UK). The probe tip is made of industrial ruby of a sphere with 1mm diameter and it has been calibrated before the digitisation process. The digitisation procedure was defined as such: the sampling interval in both X and Y directions was 0.1mm, and scanning speed was 200 points per minutes, and scanning deflection is 0.5mg. The data cloud was interpolated as a surface by using a 3-D free form surface analysis software - Cloud (UCL, UK). The degree of taper of the prepared tooth was determined in a bucco-lingual (BL) plane and mesio-distal (MD) plane.



Fig. 4. Incise dental scanner (Renishaw, UK)

All results were recorded and data were analysed by means of a one-way analysis of variance (ANOVA) and student t tests for any significant differences.

## 2.4 Results

The results of this experiment were published by Rafeek et al., (2010). There were no undercuts detected on any of the laboratory specimens as compared to the clinical specimens where 12.5% were undercut. The overall mean taper for the laboratory specimens was 22.5° (BL) and 14.5° (MD). The overall mean taper for the clinical specimens was 24.2° (BL) and 19.6° (MD). Table 1 shows the percent of specimens falling within the ideal taper range of 4° - 14°.

Specimens	BL taper between 4° - 14°.	MD taper between 4° - 14°.
Laboratory	28%	54%
Clinical	28%	25%

Table 1. Shows the percentage of tapers falling within the ideal range of 4° - 14°.

Although the mean tapers achieved in the laboratory compared to the clinic were lower, there was no significant difference detected between the mean tapers achieved ( $p > 0.05$ ) except for the MD taper of posterior teeth ( $p < 0.05$ ).

## 2.5 Discussion

The overall mean tapers achieved were higher than the recommended range of 4°-14° however the results were similar to many other clinical studies involving dental students. Table 2 summarises the comparisons of mean taper achieved in this study versus those measured with traditional devices.



CLINICAL SPECIMENS (Posteriors)	BL mean taper	MD mean taper
Rafeek et al., 2010 (Trinidad dental school)	16.8°	22.4°
Noonan & Goldfogel, 1991 (US dental school)	19.5°	18.5°
Sato et al., 1998 (Japanese dental school)	18.8°	19.2°
Patel et al., 2005 (UK dental school)	14.7°	16.3°
LABORATORY SPECIMENS (Posteriors)	BL mean taper	MD mean taper
Rafeek et al., 2010 (Trinidad dental school)	18.2°	14.2°
Ayad et al., 2005 Egyptian, US and Saudi Arabian dental schools	15°-20°	14-°20°

Table 2. Comparison of mean bucco-lingual (BL) and mesio-distal (MD) tapers of Clinical and Laboratory specimens with this study and studies utilising traditional methodology.

The comparisons of these data demonstrate that dental students in the Trinidad dental school have produced tapers that are comparable to dental students in many other dental schools around the world (Noonan & Goldfogel, 1991; Sato et al., 1998; Patel et al., 2005; Ayad et al., 2005; Mack, 1980). The dental students in the metrology study not only achieved mean tapers comparable to other students but also to dentists in a US study (Nordlander et al., 1988).

Because of the varying methods used to measure taper in the past, one must compare data with caution, however, it would appear from this work that measurements made by 3-D coordinate metrology are in close range of several studies previously undertaken. The traditional methodologies however do involve forming a two dimensional profile from the three dimensional die replica. This in itself can introduce error. This 3-D coordinate metrology is unique in that it is truly 3-D free-form surface digitisation, rather than 2-D profiling. The software is also able to use the image in a particular requested plane, therefore the taper angles can be calculated in the same plane for each of the samples. This reduces the possible error in angle calculation due to the samples being misaligned. It would be useful to measure the taper of the dies using traditional devices and then measure the taper of the same dies using CMM and compare the data to determine how closely they measure up.

## 2.6 Standards around the area

Although dimensional measurements have been widely applied in dental research, for some of the measurements, such as taper angle measurement, a consensus has been demonstrated (Noonan & Goldfogel, 1991; Sato et al., 1998; Patel et al., 2005; Ayad et al., 2005; Mack, 1980; Rafeek et al., 2006, 2010). In other areas, such as tooth wear measurements (Zhou and

Zheng, 2008; Pintado MR et al, 2000; Mitchell and Chadwick, 1998) this has not been the case. There have been some 640 papers published in tooth wear measurement over the period of 1960 to 2007, and the measurements within these papers are wide ranging using a clinical tooth wear index, a modified index, reduced depth of tooth surface in the laboratory, reduced volume and changes in percentage of the area or volume. Moreover the accuracy of these measurements were reported by accuracy, precision, inaccuracy, error and uncertainty, which demonstrated a lack of a standard in the measurement strategy, therefore it is often difficult to compare the results directly between the laboratories and published papers. The ISO is in the process of establishing and organising a committee for the dental computer assisted design-computer assisted manufacture (CAD/CAM) and its related issues so that more formal standards and protocols may be set up to improve dimensional measurements in dentistry. However, this standard will not assist in the comparison of research data, where the inaccuracies listed above are inherent. Employment of instruments with high levels of accuracy such as the CMM as a 'norm' in this type of clinical research would enable data from different laboratories to be compared at the appropriate level and power and thus enable more meaningful conclusions to be made.

### 3. Conclusion

The 3-D coordinate metrology described study shows that the taper achieved by dental students of the Trinidad dental school when preparing teeth for crowns was comparable to those achieved by other students in the US, UK, Europe, Japan, Egypt and Saudi Arabia. The study also demonstrates a method for measurement of free form oral surfaces that has enormous potential. As 'digital' dentistry develops, there is a need for worldwide standards and quality control and 3-D coordinate metrology is an extremely useful tool for this at both preclinical and clinical stages. In addition, the use of 3-D coordinate metrology in dental education, where the student is able to get 3-D visual 'feedback' of preparations in the laboratory is a valuable experience in the 'rehearsal' scenario. Lessons learned in this scenario can then be transferred to the clinic and the treatment of patients, where improved preparation quality will inevitably lead to fewer crowns requiring to be remade with all the attendant stress for both patient and clinician.

3-D coordinate metrology also has the ability to be used in a variety of areas of dental research, in particular in relation to tooth wear which is becoming an increasing global problem.

### 4. References

- Allred H (1977). A series of monographs on the assessment of the quality of dental care. The London Hospital Medical College.
- Ayad MF, Maghrabi AA, Rosenstiel SF (2005). Assessment of convergence angles of tooth preparations for complete crowns among dental students. *Journal of Dentistry*, 33: pp 633-638
- Aung SC, Ngim RCK & Lee ST (1995). Evaluation of the laser scanner as a surface measuring tool and its accuracy compared with direct facial anthropometric measurements. *British Journal of Plastic Surgery*, Vol.48, pp 551-558.

- Bosch J A(Ed.) (1995). *Co-ordinate measuring machines and systems*, Marcel Dekker, INC, ISBN 0-8247-9581-4, New York, USA.
- Bradley C & Vickers GW (1992). Automated rapid prototyping utilizing laser scanning and free-form machining, *Ann CIRP*, pp 437-440.
- Chang M & Lin PP (1999). On-line free form surface measurement via a fuzzy-logic controlled scanning probe. *International Journal of Machine tools and manufacture*, Vol.39, pp 539-552.
- Chadwick RG (1989). A review: the assessment of the durability of composite resin restorative materials in vivo. *Clinical Materials*. 4, pp 241-253.
- Cherukara GP, Seymour KG, Samarawickrama DYD, Zou LF (2002). A study into the variations in the labial reduction of teeth prepared to receive porcelain veneers – a comparison of three clinical techniques. *British Dental Journal* 192: pp 401 – 404.
- Chow JG (1997). Reproducing aircraft structural components using laser scanning. *International Journal of Advanced Manufacturing Technology*, Vol.13, pp 723-728.
- Dalton G (1998). Reverse engineering using laser metrology. *Sensor review*, Vol.18, No.2, pp 92-96.
- Dong WP, Mainsah E, Stout KJ & Sullivan PJ (1994). Three-dimensional surface topography: Review of present and future trends, Part II of Three-dimensional surface topography: measurement, interpretation and applications (Ed, Stout KJ), Penton Press, ISBN 1 85718 004 6.
- Eames WB, O'Neal SJ, Monteiro J, Miller C, Roan JD & Cohen KS. Techniques to improve the seating of castings (1978). *Journal of the American Dental Association*, 96: pp 432-437.
- Fan KC (1997). A non-contact automatic measurement for free-form surface profiles, *Computer integrated manufacturing systems*, Vol.10, No.4, pp 277-285.
- Jarvis RA (1983). A perspective on range finding techniques for computer vision. *IEEE Pattern Anal Mach Intell*, Vol.5, No.2, pp 122-139.
- Johnson GH, Craig RG (1985). Accuracy of four types of rubber impression materials compared with time of pour and a repeat pour of models. *Journal of Prosthetic Dentistry*, 53: pp 484-490.
- Jorgensen KD (1955). The relationship between retention and convergence angle in cemented veneer crowns. *Acta Odontologica Scandinavica*, 13, pp 35-40.
- Kabban M, Fearn J, Jovanovski V & Zou L (2001). Tooth size and morphology in twins; *International Journal of Paediatric Dentistry*, Vol.11, pp 333-339.
- Kent WA, Shillingburg HT, Duncason MG (1988). Taper of clinical preparations for cast restorations. *Quintessence International*, 19, pp 339-345.
- Lee KH, Park H & Son S (2001). A framework for laser scan planning of freeform surfaces; *International Journal of Advance Manufacturing Technology*, Vol.17, pp 171-80.
- Mack PJ (1980). A theoretical and clinical investigation into the taper achieved on crown and inlay preparations. *Journal of Oral Rehabilitation*, 7, pp 255-265.
- Milroy M J, Weir D J, Bradley C & Vickers G W (1996). Reverse engineering employing a 3-D laser scanner: a case study. *International Journal of Advanced Manufacturing technology*, Vol.12, pp 111-121.
- Mitchell H L and Chadwick R G (1998). Mathematical shape matching as a tool in tooth wear assessment – development and conduct. *Journal of Oral Rehabilitation* Vol.25, pp 921-8.

- Morris BO, Zou L, Royle M & Shelton C (2011). Wear Assessment of Metal-on-Metal Acetabular Cups Using a Coordinate Measuring Machine. *Wear* 271, pp 1086-1092.
- Noonan JE & Goldfogel MH (1991). Convergence of axial walls of full veneer crown preparations in a dental school environment. *Journal of Prosthetic Dentistry*, 66, pp 706-708.
- Nordlander J, Weir D, Stoffer W, et al. (1988). The taper of clinical preparations for fixed prosthodontics. *Journal of Prosthetic Dentistry*, 60, pp 148-151.
- Ohm H & Silness J (1978). The convergence angle in teeth prepared for artificial crowns. *Journal of Oral Rehabilitation*, 5, pp 371-375.
- Patel PB, Wildgoose DG & Winstanley RB (2005). Comparison of convergence angles achieved in posterior teeth prepared for full veneer crowns. *European Journal of Prosthodontics and Restorative Dentistry*, 13, pp 100-104.
- Pintado MR, DeLong R, Ko CC, Sakaguchi RL & Douglas WH (2000); Correlation of noncarious cervical lesion size and occlusal wear in a single adult over a 14-year time span; *The Journal of Prosthetic Dentistry*, Vol.84, No.4 pp 436-43.
- Rafeek RN, Marchan SM, Seymour, Zou LF & Samarawickrama DYD (2006). Abutment taper of full cast crown preparations by dental students in the UWI School of Dentistry. *European Journal of Prosthodontics and Restorative Dentistry*, 2, pp 63-66.
- Rafeek RN, Smith WA, Seymour KG, Zou LF & Samarawickrama DYD (2010). Taper of full-veneer crown preparations by dental students at the University of the West Indies. *Journal of Prosthodontics*, Oct, 19(7), pp 580-585.
- Robinson PB & Lee JW (2001). The use of real time video magnification for the pre-clinical teaching of crown preparations. *British Dental Journal*, 190(9), pp 506-510.
- Rosenstiel E (1975). The taper of inlay and crown preparations. *British Dental Journal*, 139, pp 436-438.
- Sato T, Mutawa N, Okada D & Hasegawa S (1998). A clinical study on abutment taper and height of full cast crown preparations. *Journal of Medical and Dental Sciences*, 45, pp 205-210.
- Seymour KG, Zou LF, Samarawickrama DYD & Lynch E (1996). Assessment of shoulder dimensions and angles of porcelain bonded to metal crown preparations. *Journal of Prosthetic Dentistry*, 75, pp 406 - 411.
- Shahrom AW, Vanezis P, Chapman RC, Gonzales A, Blenkinsop C & Rossi M L (1996). Techniques in facial identification: computer aided facial reconstruction using a laser scanner and video superimposition. *International Journal of Legal Medicine*, Vol.108, pp 194-200.
- Shillingburg H, Hobo S, Whitsett LD, et al. (1997). Fundamentals of fixed prosthodontics. Chicago; Quintessence., pp119-138.
- Smith CT, Gary JJ, Conkin JE, et al. (1999). Effective taper criterion for the full veneer crown preparation in preclinical prosthodontics. *Journal of Prosthodontics*, 8, (3), pp 196-200.
- Stout K J (1994). Three-dimensional surface topography: measurement, interpretation and applications; Penton press; ISBN 1 85718 004 6.
- Summerhays K D, Henke R P, Baldwin J M, Cassou RM & Brown CW (2002). Optimizing discrete point sample patterns and measurement data analysis on internal cylindrical surfaces with systematic form deviations. *Precision engineering, Journal of the International Societies for precision engineering and nanotechnology*, vol.26 pp 105-121.

- Theocharopoulos A, Zou L, Hill R & Cattell M (2010). Wear quantification of human enamel and dental glass-ceramics using white light profilometry. *Wear* 269, pp 930-936.
- Weed RM, Suddick RP, Kleffner JH (1984). Taper of clinical and typodont crowns prepared by dental students. (Abstract) *Journal of Dental Research*, 63, pp 286.
- Win P, Van Vliet W & Schellekens P (1998). Development of a fast mechanical probe for coordinate measuring machines. *Precision engineering, Journal of the International Societies for precision engineering and nanotechnology*, 22, pp 141-152.
- Zhou Z R & Zheng J (2008); Tribology of dental materials: A Review. *Journal of Physics D Applied physics*, Vol.41, pp 113001-23.
- Zou L, Cherukara G, Hao P, Seymour K & Samarawickrama DYD (2009). Geometrics of tooth wear. *Wear* Vol.266, pp 205-208.
- Zou L, Samarawickrama DYD, Jovanovski V & Shelton J (2001). Measurements of sequential impressions of acetabula cups from a total hip joint replacement using a non-contact measurement system. *International Journal of Machine tools & manufacture*, Vol.41, pp 2023-2030.