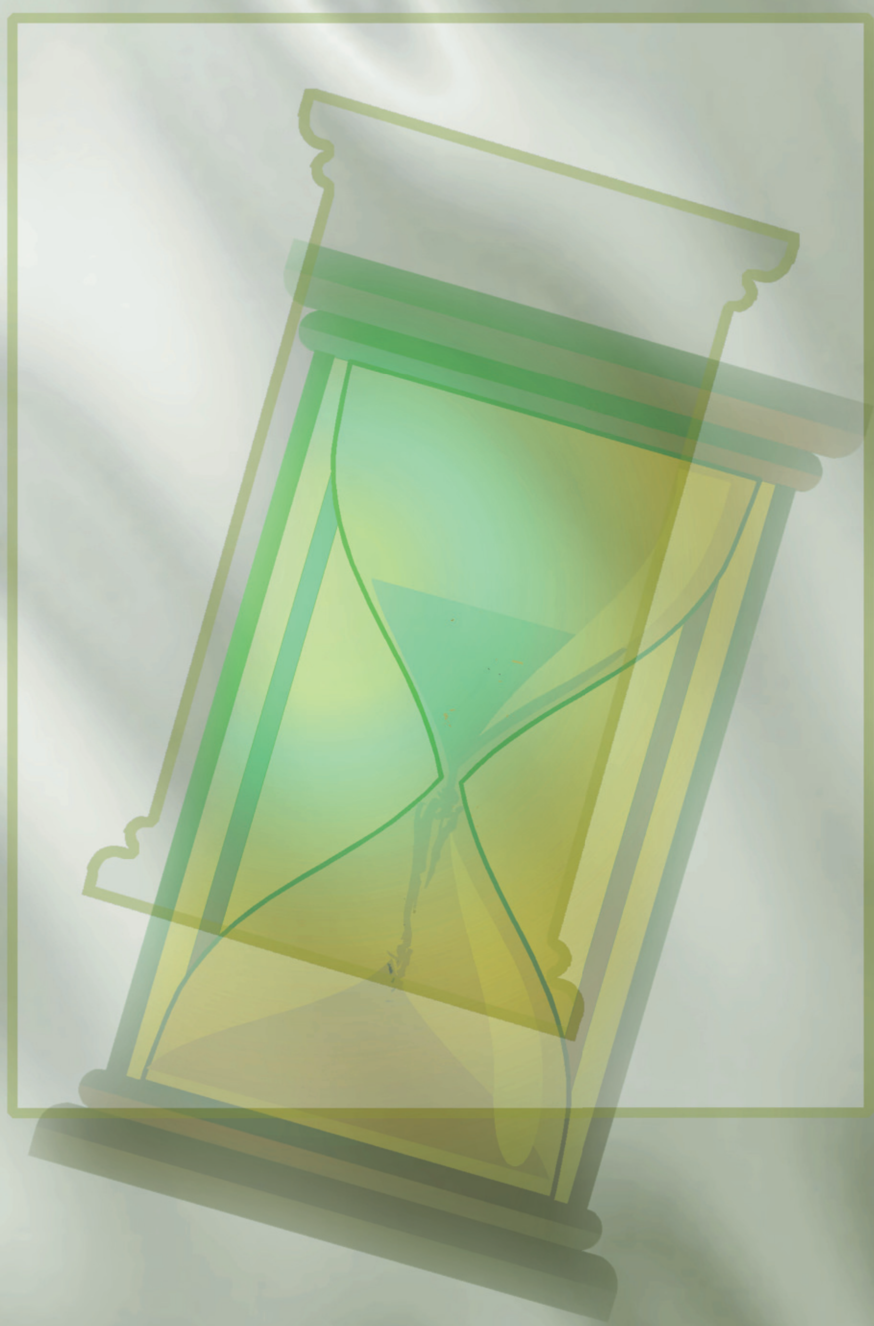
INTERNATIONAL TELECOMMUNICATION UNION

HANDBOOK
SELECTION AND USE OF PRECISE FREQUENCY AND TIME SYSTEMS



1997

RADIOCOMMUNICATION BUREAU

THE RADIOCOMMUNICATION SECTOR OF THE ITU

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Contact address for inquiries about radiocommunication matters:

ITU
Radiocommunication Bureau
Place des Nations
CH-1211 Geneva 20
Switzerland

Telephone	+41 22 730 5800
Fax	+41 22 730 5785
Internet	brmail@itu.int
X.400	S=brmail; P=itu; A=400net; C=ch

Contact address for orders of ITU publications:

ITU
Sales and Marketing Service
Place des Nations
CH-1211 Geneva 20
Switzerland

Telephone	+41 22 730 6141 English
Telephone	+41 22 730 6142 French
Telephone	+41 22 730 6143 Spanish
Fax	+41 22 730 5194
Telex	421 000 uit ch
Telegram	ITU GENEVE
Internet	sales@itu.int
X.400	S=sales; P=itu; A=400net; C=ch

INTERNATIONAL TELECOMMUNICATION UNION

HANDBOOK

SELECTION AND USE OF PRECISE FREQUENCY AND TIME SYSTEMS

1997

RADIOCOMMUNICATION BUREAU

PREFACE

This Handbook on the Selection and Use of Precise Frequency and Time Systems has been developed by the Group of the Experts in the field of frequency standards and time signals of Radiocommunication Study Group 7 – Science Services under the Chairmanship of Mr. R. Sydnor (Editor, USA) and Mr. David W. Allan (Assistant Editor, USA).

This Handbook consists of 10 Chapters which describe basic concepts, frequency and time sources, measurement techniques, characteristics of various frequency standards, operational experience, problems and future prospects. The contents include detailed explanations and many references that can be consulted for additional details.

The technical content of this Handbook is intended for the use by administrations of both developing and developed countries and by the Radiocommunication Bureau. The Handbook will also be useful to engineers of the scientific and industrial organizations.

Robert W. Jones

Director, Radiocommunication Bureau

Introduction by Radiocommunication Study Group 7 Chairman

The Radiocommunication Study Group for the Science Services (SG-7) was created through a structural reorganisation in 1990 at the Dusseldorf CCIR Plenary Assembly. At that time, the Space Research and Radio Astronomy Study Group (SG-2) was consolidated with the Time and Frequency Standards Study Group (SG-7) to form a new Study Group 7 on Science Services.

Many of the activities in the Science Services SG are associated with advancing the state of the art in the use of the radio spectrum to achieve scientific objectives. In this regard, the time and frequency standards community has long been associated with the International Telecommunication Union with the express purpose of developing Recommendations for the use of the radio spectrum to facilitate the dissemination of precise time references and for standardising the methods for this dissemination. An essential corollary is the specification of precise frequency standards and the techniques for their implementation.

While the development of Recommendations was, and continues to be, the principal focus of the Study Group activities, it has become clear that the experts who work on these matters in the Study Group have much basic information to offer to their scientific and lay colleagues who depend on precise time and frequency data for implementing a variety of communications techniques. These techniques include generation of official time for all nations, telecommunications, navigation (including collision avoidance), power systems, position determination and surveying, avionics, transportation systems, space exploration, astronomy and astrometry (especially millisecond pulsar measurements), earthquake monitoring, and all national standards laboratories.

Thus it was decided to prepare and publish this Handbook so that all users of these standards could more completely understand precise time and frequency sources and systems in order to better design and apply these powerful tools.

As Chairman of Study Group 7, it is my honour and pleasure to present this Handbook to the community of users of precise time and frequency standards who will, I am sure, find it an invaluable reference tool in their own work.

H. G. Kimball

Geneva 1996

Acknowledgements

We should like to thank the following administrations and organizations for their kind support and participation on the part of their experts:

France, Germany (Federal Republic of), Italy, Switzerland (Confederation of), United States of America, Hewlett-Packard Laboratories (USA), Istituto Elettrotecnico Nazionale G. Ferrar (Italy), Jet Propulsion Laboratory (USA), Laboratoire de L'Horloge Atomique (France), National Institute of Standards and Technology (USA), Observatoire de Neuchâtel (Switzerland), Observatoire de Paris (France), Physikalische-Technische Bundesanstalt (Germany), Polytechnico di Torino (Italy) and Bureau international des poids et mesures (BIPM),

and to acknowledge the contributions and the helpful guidance and discussions with the following personalities:

Chapter 1	Claude Audoin	Laboratoire de L'Horloge Atomique, France;
Chapter 2A	Andreas Bauch	Physikalische-Technische Bundesanstalt, Germany;
Chapter 2B	Roger Beehler	National Institute of Standards and Technology, USA;
Chapter 3	Laurent-Guy Bernier	Observatoire de Neuchâtel, Switzerland;
Chapter 4	Fred Walls	National Institute of Standards and Technology, USA;
Chapter 5	Richard L. Sydnor	Jet Propulsion Laboratory, USA;
Chapter 6	Claudine Thomas	Bureau international des poids et mesures, France;
Chapter 7	Sigfrido Leschiutta	Polytechnic di Totino, Italy;
	Franco Cordara	Istituto Elettrotecnico Nazionale G. Ferrar, Italy;
Chapter 8	Michel Granveaud	Observatoire de Paris, France;
Chapter 9	Leonard Cutler	Hewlett-Packard Laboratories, USA;
Chapter 10	Donald Sullivan	National Institute of Standards and Technology, USA;

and also to thank Messrs. R. L. Sydnor and D.W. Allan for their editorial work.

Table of Contents

Introduction to the Handbook

Over the last few decades atomic clocks have moved from laboratory novelties to large scale usage. The improvement in quartz oscillator technology and in satellite timing systems has augmented the decades of improvement in atomic clocks. Navigation, communication and power systems have all benefited greatly from these improvements. Precise timing has moved from a novelty to a necessity, and we are seeing many other applications benefit as well. One of the reasons being that frequency is not only the most accurate measurement known to man, but it is also the most cost effective of the metrology units. Hence, many applications depend on precise timing elements.

This Handbook has been prepared in response to this rapidly increasing and more generalised user community. The goal is to provide help and guidance to the seasoned as well as the new and future users of precise timing to best meet their needs. The Handbook deals with source selection and proper utilisation.

Because this Handbook addresses a much broader audience than just the time and frequency community, an effort has been made to explain the nomenclature and symbols used. A Glossary is included in the front. Specifically, Chapter 3 on characterising time and frequency sources has been written using communication engineering terminology. The Handbook as a whole conforms with IEEE standards for nomenclature with the exception of Chapter 3, and in Chapter 3 there is a conversion table so that those with communication engineering background can know the appropriate IEEE designation for terms of reference and vice versa. Chapter 10, the summary chapter, ties together the importance of each of the chapters and is probably worth reading first as it provides both a good beginning and ending for this work.

A significant effort has been made to maintain overall consistency and continuity in the Handbook. Chapter 1 and Chapter 2 complement each other in providing the physics and the sources of time and frequency standards and systems. Chapter 4 (which is more experimental) builds on Chapter 3 (which is more theoretical), and together provide the basis for characterising these time and frequency sources and systems. Chapter 5 summarises performance, including environmental effects, using the information in the first four chapters. Chapter 6 shows the power and importance of clock ensembles using the tools from previous chapters – plus some very important ensembling concepts. This is a very important chapter for timing centres to understand (new or old). Chapter 7 shows the myriad of applications existing now and some anticipated applications. Chapter 8 is a “things to watch out for chapter” written so that others don’t repeat the mistakes of the past and so the reader can learn from the experience of the experts in the field. Chapter 9 is very helpful for the planner as the future is projected. Again, Chapter 10 is a highlighted summary of each of the chapters and also showing in more detail how the Handbook can be used and how the chapters interrelate.

The authors of the chapters have been picked from among the best in the world in this field of endeavour. The authors and the editors have dedicated their efforts toward the Handbook hoping it will be a basic and useful reference resource for many years to come.

One very important conclusion in the selection and use of precise frequency and time sources and systems is that the ability to obtain high accuracy at low cost has improved dramatically over the last decade. This has been mainly due to satellite techniques – principally GPS. Properly configured, a couple of atomic clocks and a GPS receiver can make a very accurate and reliable timing system. We anticipate that as satellite techniques mature, time and frequency sources and systems will naturally follow. Terrestrial time and frequency distribution techniques will probably continue to fade in importance.

Glossary

1 Introduction

The list of terms below is a glossary for the users of standard-frequency and time-signal services. It is taken from Recommendation ITU-R TF.686, the International Vocabulary of basic and general terms in Metrology (VIM), published by ISO, and other terms in general usage in the field of frequency and time. Precise time measurements may often be affected by relativity effects. The terms and definitions below do not in all cases imply incorporation of, or indicate the need for, the consideration of these effects. Two types of terms are presented; those typically used within the standard-frequency and time-signal services and those of more general use, but specifically relevant to this field. For the latter, an attempt has been made to provide substantial agreement with the definitions contained in the International Electrotechnical Vocabulary (IEV). The equivalence of the terms are given in French and Spanish (terms printed in italics). Terms taken from Recommendation ITU-R TF.686 are not marked, terms from the ITU Radio Regulations and the ISO publications are indicated as “RR” and “ISO”, respectively. Additional terms may be found in the National Institute of Standards and Technology Technical Note 1297, 1994, “Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results.”

2 Definitions

The definitions follow in the alphabetical order.

Accuracy; *Exactitude*; *Exactitud*

The degree of conformity of a measured or calculated value to its definition (see “uncertainty”).

(ISO) Accuracy is generally characterised by the overall uncertainty of a measured value.

Ageing; *Vieillissement*; *Envejecimiento*

The systematic change in frequency with time due to internal changes in the oscillator.

Note – It is the frequency change with time when factors external to the oscillator (environment, power supply, etc.) are kept constant.

Atomic Time Scale; *Échelle De Temps Atomique*; *Escala De Tiempo Atómico*

A time scale based on atomic or molecular resonance phenomena.

Calibration*; *Étalonnage*; *Calibración*

The process of identifying and measuring offsets in instruments and/or procedures.

Note – In many cases, e.g. in a frequency generator, the calibration is related to the stability of the device and therefore its result is a function of time and of the averaging time.

(ISO) The process of identifying and measuring deviations between the indicated value and the value of a reference standard used as the test object.

Clock; *Horloge*; *Reloj*

A device for time measurement and/or time display.

* This definition differs from those in the IEV, but ITU-R is of the opinion that they are more appropriate for the standard-frequency and time-signal service.

Clock time difference; *Différence entre temps d'horloge*; *Diferencia de tiempo de reloj*

The difference between the readings of two clocks at the same instant.

Note – In order to avoid confusion in sign, algebraic quantities should be given, applying the following convention. At a time T of a reference time scale, let a denote the reading of a time scale A, and b the reading of a time scale B; the time scale difference is expressed by $A - B = a - b$ at the instant T. The same convention applies to the case where A and B are clocks.

Coherence of frequency; *Cohérence de fréquence*; *Coherencia de frecuencia*

See "coherence of phase".

Coherence of phase; *Cohérence de phase*; *Coherencia de fase*

Exists if two periodical signals of frequency M and N resume the same phase difference after M cycles of the first and N cycles of the second, M/N being a rational number, obtained through multiplication and/or division from the same fundamental frequency.

Coordinated clock; *Horloge coordonnée*; *Reloj coordinado*

A clock synchronized within stated limits to a reference clock which is spatially separated. (See also Report 439 (Dusseldorf, 1990), which deals with the concept of coordinate time.)

Coordinate time; *Temps-coordonnée*; *Tiempo-coordenada*

The concept of time in a specific coordinate frame, valid over a spatial region with varying gravitational potential.

Note – If a time scale is realized according to the coordinate time concept, it is called a coordinate time scale.

Example: TAI is a coordinate time scale in an earth-based reference frame with the SI second as realised on the rotating geoid as the scale unit.

Coordinated time scale; *Échelle de temps coordonnée*; *Escala de tiempo coordinada*

A time scale synchronized within stated limits to a reference time scale.

Coordinated Universal Time (UTC); *Temps universel coordonné*; *Tiempo Universal Coordinado*

The time scale, maintained by the BIPM and the International Earth Rotation Service (IERS), which forms the basis of a coordinated dissemination of standard frequencies and time signals (see Recommendation ITU-R TF. 460).

It corresponds exactly in rate with TAI, but differs from it by an integral number of seconds. The UTC scale is adjusted by the insertion or deletion of seconds (positive or negative leap seconds) to ensure approximate agreement with UT1.

Date; *Date*; *Fecha*

The reading of a specified time scale.

Note – The date can be conventionally expressed in years, months, days, hours, minutes, seconds and fractions thereof. Also, "Julian Date" (JD) and "Modified Julian Date" (MJD) are useful dating measures (see "Julian Date" and "Modified Julian Date").

Drift (implying frequency drift); *Dérive*; *Deriva*

The systematic change in frequency with time of an oscillator.

Note – Drift is due to ageing plus changes in the environment and other factors external to the oscillator (see ageing).

DUT1; *DUT1*; *DUT1*

The value of the predicted difference UT1 – UTC, as disseminated with the time signals. DUT1 may be regarded as a correction to be added to UTC to obtain a better approximation to UT1. The values of DUT1 are given by the IERS in integral multiples of 0.1 s (see Universal Time).

Error * ; *Erreur*; *Error*

The difference of a value from its assumed correct value.

(ISO) Result of a measurement minus a true value

Frequency*; *Fréquence*; *Frecuencia*

If T is the period of a repetitive phenomenon, then the frequency $f = 1/T$. In SI units the period is expressed in seconds, and the frequency is expressed in hertz.

Frequency difference; *Différence de fréquence*; *Diferencia de frecuencia*

The algebraic difference between two frequency values.

Frequency drift*; *Dérive de fréquence*; *Deriva de frecuencia*

See "drift" and "ageing".

Frequency instability; *Instabilité de fréquence*; *Inestabilidad de frecuencia*

The spontaneous and/or environmentally caused frequency change within a given time interval.

Note – Generally one distinguishes between systematic effects such as frequency drift effects and stochastic frequency fluctuations. Special variances have been developed for the characterization of these fluctuations. Systematic instabilities may be caused by radiation, pressure, temperature, humidity, etc. It is typically dependent on the measurement system bandwidth and/or on the sample time or integration time. Random or stochastic instabilities are typically characterized in the time-domain and/or frequency-domain (Recommendation ITU-R TF.538).

In many contexts the expression "stability" instead of "instability" is used. This usage is acceptable.

Frequency offset; *Décalage de fréquence*; *Separación de frecuencia*

The systematic frequency difference between the realized value and the nominal frequency value.

Frequency shift; *Déplacement de fréquence*; *Desplazamiento de frecuencia*

An intentional frequency change.

Frequency stability; *Stabilité de fréquence*; *Estabilidad de frecuencia*

See "frequency instability".

Frequency standard; *Étalon de fréquence*; *Patrón de frecuencia*

A generator, the output of which is used as a frequency reference.

Note – See "Primary frequency standard" and "Secondary frequency standard".

Instant; *Instant*; *Instante*

A point in time.

---

* This definition differs from those in the IEV, but ITU-R is of the opinion that they are more appropriate for the standard-frequency and time-signal service.

International Atomic Time (TAI); *Temps atomique international*; *Tiempo Atómico Internacional*

The time scale established by the Bureau international des poids et mesures (BIPM) on the basis of data from atomic clocks operating in several establishments conforming to the definition of the second, the unit of time of the International System of Units (SI).

Julian Date (JD); *Date julienne (DJ)*; *Fecha Juliana (FJ)*

The Julian Day Number followed by the fraction of the day elapsed since the preceding noon (12h00 UT).

Example: The date 1900 January 0.5 d UT corresponds to JD = 2 415 020.0.

Note – The Julian Date is conventionally referred to UT1, but may be used in other contexts, if so stated.

Julian day number; *Numéro de jour julien*; *Número de día juliano*

The number of a specific day from a continuous day count having an initial origin of 12h00 UT on 1 January 4713 BC, Julian proleptic Calendar (start of Julian Day zero).

Example: The day extending from 1900 January 0.5 d UT to 1900 January 1.5 d UT has the number 2 415 020.

Leap second; *Seconde intercalaire*; *Segundo intercalar*

An intentional time step of one second used to adjust UTC to ensure approximate agreement with UT1. An inserted second is called positive leap second and an omitted second is called negative leap second (see Recommendation ITU-R TF.460).

Modified Julian Date (MJD); *Date julienne modifiée*; *Fecha Modificada del Calendario Juliano*

Julian Date less 2 400 000.5 days (see Recommendation 457).

Modified Julian Day

Integer part of Modified Julian Date.

Nominal value * ; *Valeur nominale*; *Valor nominal*

A specified or intended value independent of any uncertainty in its realization.

Note – In a device that realizes a physical quantity, it is the specified value of such a quantity. It is an ideal value and thus it is free from tolerance.

Normalized value; *Valeur normée*; *Valor normalizado*

The ratio of a value to its nominal value.

Note 1 – This definition can be used in connection with: frequency, frequency deviation, frequency difference, frequency drift, frequency offset, etc.

Note 2 – In place of the term “normalized”, the term “relative” is acceptable but the term “fractional” is to be avoided.

Offset ***;** *Décalage*; *Separación*

The systematic difference between the realized value and the nominal value. (See also “Normalized offset”.)

Phase; *Phase*; *Fase*

Generally in a periodic phenomenon, analytically described by a function of time (or space), the phase is any possible and distinguishable state of the phenomenon itself.

---

* This definition differs from those in the IEV, but ITU-R is of the opinion that they are more appropriate for the standard-frequency and time-signal service.

It can be identified through the time of its occurrence, elapsed from a specified reference, to be called correctly “phase time” (frequently abbreviated to “phase”). Particularly, if the phenomenon is sinusoidal, the phase can be identified either by the angle or by the time, both measured from an assigned reference, depending on the dimensions assigned to the reference period (namely 2p or T).

In the standard-frequency and time-signal service, phase-time differences are mainly considered, i.e. time differences between two identified phases of the same phenomenon or of two different phenomena.

Phase shift; *Déphasage*; *Desplazamiento de fase*

An intentional change in phase from a reference.

Phase deviation; *Décalage de phase*; *Desviación de fase*

The difference of the phase from a reference.

Precision; *Précision*; *Precisión*

The degree of mutual agreement among a series of individual measurements; often, but not necessarily, expressed by the standard deviation.

Primary frequency standard; *Étalon primaire de fréquence*; *Patrón primario de frecuencia*

A frequency standard whose frequency corresponds to the adopted definition of the second, with its specified accuracy achieved without external calibration of the device.

Note – The second is defined as follows:

> “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 caesium atom-133.” (XIII[e] Conférence générale des poids et mesures, 1967.)

Primary clock; *Horloge primaire*; *Reloj primario*

A time standard which operates without external calibration (see “time standard”).

Proper time; *Temps propre*; *Tiempo propio*

The local time, as indicated by an ideal clock, in a relativistic sense (see Report 439).

Note – This is distinguished from a coordinate time, which involves theory and computations.

If a time scale is realized according to the proper time concept, it is called a proper time scale.

Examples:

a) for proper time: the second is defined in the proper time of the caesium atom;

b) for proper time scale: a time scale produced in a laboratory, not transmitted outside the laboratory.

Repeatability, *Répétabilité,*

(ISO) Closeness of agreement between the results of successive measurements of the same measured carried out under the same conditions of measurement:

- a) with respect to a single device when specified parameters are independently adjusted under stated conditions of use, it is the standard deviation of the values produced by this device. It should better be called: repeatability of indication of one device with readjustment, resettability.
- b) with respect to a single device put into operation repeatedly without readjustment, it is the standard deviation of the values produced by this device. It should better be called: repeatability of indication of one device without readjustment.
- c) with respect to a set of independent devices of the same design, it is the standard deviation of the values produced by these devices. It should better be called: repeatability of indication of different devices of the same design used under the same conditions.

Reproducibility; *Reproductibilité*; *Reproductibilidad*

a) With respect to a set of independent devices of the same design, it is the ability of these devices, to produce the same value.

b) With respect to a single device, put into operation repeatedly without adjustments, it is the ability to produce the same value.

Note – A usual measure of the lack of reproducibility is the standard deviation.

Resettability [*]**;** *Fidélité*; *Reposicionabilidad*

It is the ability of a device to produce the same value when specified parameters are independently adjusted under stated condition of use. This term should be used instead of the "repeatability", considered as not pertinent to frequency generators.

Note – A usual measure of the lack of resettability is the standard deviation.

Secondary frequency standard; *Étalon secondaire de fréquence*; *Patrón secundario de frecuencia*

A frequency standard which requires external calibration.

Standard frequency; *Fréquence étalon*; *Frecuencia patrón*

A frequency with a known relationship to a frequency standard.

Note 1 – The term standard frequency is often used for the signal whose frequency is a standard frequency.

Note 2 – The term standard frequency is often used for a frequency that is one of a set of ITU-R agreed values, i.e., 1 MHz, 5 MHz, etc.

Standard frequency and time-signal station; *Station de fréquence étalon et/ou de signaux horaires*; *Estación de frecuencias patrón y/o de señales horarias*

A station which provides a standard-frequency and/or time-signal emissions.

(RR) A station in the standard frequency and time signal service.

Standard-frequency emission; *Émission de fréquences étalon*; *Emisión de frecuencias patrón*

An emission which disseminates a standard frequency at regular intervals with a specified frequency accuracy.

Note – Recommendation ITU-R TF.460 recommends a normalized frequency deviation of less than 1×10^{-10}.

Standard Frequency and Time Signal Service, *Service des fréquences étalon et des signaux horaires, Servicio de frecuencias patrón de señales horarias*

(RR) A radiocommunication service for scientific, technical and other purposes, providing the transmission of specified frequencies, time signals, or both, of stated high precision, intended for general reception.

Standard Frequency and Time Signal-Satellite Service, *Service des fréquences étalon et des signaux horaires par satellite, Servicio de frecuencias patrón de señales horarias por satélite*

(RR) A radiocommunication service using space stations on earth satellites for the same purpose as those of the standard frequency and time signal service.

Standard-time-signal emission; *Émission de signaux horaires*; *Emisión de señales horarias*

An emission which disseminates a sequence of time signals at regular intervals with a specified accuracy.

---

[*] This definition differs from those in the IEV, but ITU-R is of the opinion that they are more appropriate for the standard-frequency and time-signal service.

Note – Recommendation ITU-R TF.460 recommends standard time-signals to be emitted within 1 ms with reference to UTC and to contain DUT1 information in a specified code.

Synchronism; *Synchronisme*; *Sincronismo*

See “time scales in synchronism”.

Time; *Temps*; *Tiempo*

Note – In English “time” is used to specify an instant (time of day) or as a measure of time interval.

Time comparison; *Comparaison de temps*; *Comparación de tiempos*

The determination of a time scale difference.

Time code; *Code horaire*; *Código horario*

An information format used to convey time information.

Time interval; *Intervalle de temps*; *Intervalo de tiempo*

The duration between two instants.

Time marker; *Repère de temps*; *Marca de tiempo*

A reference signal enabling the assignment of dates on a time scale.

Time scale; *Échelle de temps*; *Escala de tiempo*

A system of unambiguous ordering of events.

Time scale difference; *Différence entre échelles de temps*; *Diferencia entre escalas de tiempo*

The difference between the readings of two time scales at the same instant.

Note – In order to avoid confusion in sign, algebraic quantities should be given, applying the following convention. At a time T of a reference time scale, let a denote the reading of a time scale A, and b the reading of a time scale B; the time scale difference is expressed by $A - B = a - b$ at the instant T. The same convention applies to the case where A and B are clocks.

Time scales in synchronism; *Échelles de temps en synchronisme*; *Escalas de tiempo en sincronismo*

Two time scales are in synchronism, when they assign the same date to an instant.

Note – If the time scales are produced in spatially separated locations, the propagation time of transmitted time signals and relativistic effects – including the reference coordinate frame – are to be taken into account (see Report 439).

Time scale reading; *Lecture d'une échelle de temps*; *Lectura de una escala de tiempo*

The value read on a time scale at a specific instant.

Note – The reading of a time scale should be qualified by giving the time scale name (see Recommendation ITU-R TF.536).

Time scale unit; *Unité d'une échelle de temps*; *Unidad de escala de tiempo*

The defining basic time interval in a time scale.

Note – This is to be distinguished from the realized time scale unit.

Time-signal satellite service; *Service des signaux horaires par satellite*; *Servicio de señales horarias por satélite*

A radiocommunication service using earth satellites for the same purpose as those of the time-signal service.

Time standard; *Étalon de temps*; *Patrón de tiempo*

a) A device used for the realization of the time unit.

b) A continuously operating device used for the realization of a time scale in accordance with the definition of the second and with an appropriately chosen origin.

(ISO) A continuously operating device used for the realisation of a time scale.

Time step; *Saut de temps*; *Salto de tiempo*

A discontinuity in a time scale at some instant.

Note – A time step is positive (+) if the time scale reading is increased, and negative (–) if the reading is decreased at that instant.

Uncertainty; *Incertitude*; *Incertidumbre*

The limits of the confidence interval of a measured or calculated quantity.

Note – The probability of the confidence limits should be specified, preferably by the one sigma value.

(ISO) Parameter associated with the result of a measurement that characterises the dispersion of the values that could reasonably be attributed to the measurand.

Frequently it is possible to distinguish two components, the random component and the component due to systematic effects.

The random uncertainty is often expressed by the standard deviation or by a multiple of the standard deviation for repeated measurements. The component due to systematic effects is generally estimated on the basis of all available information about relevant parameters.

Universal Time (UT); *Temps universel*; *Tiempo Universal*

Universal Time (UT) is the general designation of time scales based on the rotation of the Earth. In applications in which a precision of a few tenths of a second cannot be tolerated, it is necessary to specify the form of UT such as UT1, which is directly related to the rotation of the Earth as explained in Recommendation ITU-R TF.460.

Universal Time Coordinated (UTC); *Temps universel coordonné*; *Tiempo universal coordinado*

See "Coordinated Universal Time", which is an equivalent expression.

Chapter 1

Introduction and basic concepts

Contents

1.1 Historical sketch

Quartz crystal oscillators were the first frequency and time standards that took advantage of the development of electronic technology. They rely upon the discovery of piezoelectricity by P. Curie in 1880 (Nobel prize 1903) and upon the invention of the first electronic amplifier (the triode) by Lee de Forest in 1907. Such oscillators began to be used in communication systems and in frequency and time metrology laboratories between 1920 and 1930. They were largely used during World war II and, since that time, a large amount of effort has been devoted to improving the resonator design, to optimising the behaviour of the associated electronic circuit and to understanding the physical origin of the short-term frequency instabilities, the sensitivity to external factors and the ageing processes [Besson, 1977; Filler and Vig, 1993; Gerber and Ballato, 1985; Vig, 1991; Walls et al., 1992; Walls and Ggagnepain, 1992]. In the field of interest of this Handbook, high quality quartz crystal oscillators are used as secondary frequency and time standards. Furthermore, they are present in most atomic frequency standards, where their frequency is controlled by the atomic resonance and where they are the source of the output signals.

The idea of using the resonance properties of an ensemble of isolated atoms (i.e. in a vapour) to realise a frequency standard and a clock is attributed to I.I. Rabi, in 1939 (Nobel prize 1944). The first atomic clock was built at the National Bureau of Standards (NBS) during 1948-49, using a microwave absorption line of ammonia. The most important discoveries which determine the structure of present operational cæsium, rubidium and hydrogen atomic frequency standards are that of: a) the two separated field method of magnetic resonance and the hydrogen maser by N.F. Ramsey in 1950 and 1960 respectively (Nobel prize 1989); b) the optical pumping by A. Kastler in 1950 (Nobel prize 1966); and c) the maser effect in the ammonia maser by C.H. Townes, N.G. Basov and A.M. Prokhorov in 1955 (Nobel prize 1964). Research is being pursued to develop atomic frequency standards based on a resonance of ions stored in an electric trap, following W. Paul and H.G. Dehmelt (Nobel Prize 1989) [Wineland et al., 1990].

The first cæsium beam frequency standard was built at NBS in the early 1950s, and the first cæsium atomic clock began operation at National Physics Laboratory (NPL) (England) in June 1955 [Essen and Parry, 1957]. The first commercial units were available around 1958. Soon afterwards, at the beginning of the 1960's, the production of rubidium frequency standards and of hydrogen masers started also.

The NPL cæsium beam device was used in 1958 to measure the unperturbed hyperfine separation of the cæsium atom in the ground state in terms of the Ephemeris second defined at that time. The result is included in the present definition of the unit of time which was decided in 1967 by the Thirteenth Conférence Générale des Poids et Mesures: "*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 cæsium 133 atom*" [BIPM, 1991].

Laboratory cæsium beam frequency standards have been realised by major frequency and time laboratories, for instance in Canada, in China, in France, in Germany, in Japan, in Russia and in the United States of America, with the aim of realising as accurately as possible the definition of the second. At the time of writing, the best accuracy achieved with some of these primary frequency standards is about $1 \rightarrow 2 \times 10^{-14}$ [Bauch et al., 1988; Dorenwendt et al., 1990; Lee et al., 1995].

Since their appearance, the frequency stability and the reliability of manufactured atomic frequency standards using a resonance in the cæsium, rubidium or hydrogen atom have been considerably improved. Furthermore, their volume has been reduced.

The progress of laboratory primary frequency standards and of field operable units is the result of a better understanding of their physics, of the origin of the frequency instabilities and of the residual frequency offsets. Of course, the progress of the technology of the electronic circuits which are necessary to probe the atomic transition and to control a quartz crystal oscillator have contributed to these developments. A survey of the principles and of the properties of atomic frequency standards is given in the following references: Hellwig, 1985; Vanier and Audoin, 1989; Busca et al, 1990; Cutler, 1990; Vessot, 1990; Lewis, 1991; Audoin et al, 1992; Mattison, 1992; and Riley, 1992.

Present research efforts are directed towards the application of optical pumping to cæsium beam frequency standards [Audoin, 1992], the implementation of the methods of laser cooling of ions [Wineland et al., 1990; Prestage et al., 1994], laser cooling in a frequency standard using a fountain of cæsium atoms [Clairon et al., 1995; Gibble and Chu, 1993] and the development of a cryogenic hydrogen maser [Vessot et al., 1990].

1.2 Basic principles of frequency standards

In this chapter a brief overview is given on the principle of the operation of the available frequency standards as well as the basic metrological concepts for their characterisation, including stability and accuracy. A more detailed description of the devices including the characteristics of their performance is given in Part A of Chapter 2.

1.2.1 Quartz crystal frequency standards

A quartz crystal frequency standard is an electronic oscillator in which the frequency determining element is a resonator made of a quartz crystal. The resonance frequency is mainly determined by the macroscopic properties of the bulk material. It is therefore also dependent on all environmental influences that modify these dimensions and properties and can in practice not be deduced from fundamental properties of individual atoms. Therefore, their frequency needs to be calibrated with respect to that of a more accurate frequency source. High performance units are widely used in frequency and time metrology laboratories as secondary frequency standards. Furthermore, the output signal of most atomic frequency standards is derived from a voltage controlled crystal oscillator (VCXO), its frequency being controlled by the atomic resonance.

1.2.2 Atomic frequency standards

In contrast to quartz oscillators, the reference frequency in atomic frequency standards is mainly determined by the intrinsic properties of atoms of specifically selected elements, in other words, by fundamental constants which result from the basic interactions between elementary particles.

In our present state of knowledge in Physics and Astronomy, we are allowed to postulate that the atomic properties are fixed and do not depend on space and time (within known relativistic effects). Therefore, it is possible to build and to disseminate equipment which delivers the same frequency to a number of locations, provided that a given transition in a given element is observed and that the relativistic effects (related to the altitude for instance) are taken into account. A reference frequency is then available locally in real time. It remains constant with time and equal to that of other atomic frequency standards situated at different places on the Earth and its close vicinity (within uncertainties specified in Chapters 2, Part A and 5). In the practical work we are dealing almost entirely with three types of standards which make use of the properties of the elements hydrogen (the hydrogen maser), rubidium (the rubidium gas cell frequency standard) and cæsium (the cæsium beam frequency standard). These will be discussed subsequently while future developments are dealt with in Chapter 9.

1.2.2.1 Spectroscopic properties of interest

Let us consider an atom of a given element. It has quantum levels of well defined energy. Let E_1 and E_2 be the energy of two of them, with $E_2 > E_1$. A transition between these two levels may occur under the effect of an electromagnetic radiation of frequency ν_0. Energy conservation determines the value of ν_0. This leads to Bohr's relation:

$$h\nu_0 = E_2 - E_1 \tag{1.1}$$

where h is Planck' constant. Therefore a resonance may be observed. This resonance has a width $\Delta\nu$, which is given by Heisenberg's uncertainty relation:

$$\Delta\nu\,\Delta t \geq 1 \tag{1.2}$$

For the transitions considered here, Δt is the practical duration of their observation. It is limited by various physical phenomena or processes such as transit times, relaxation, etc., and it ranges from approximately

1 ms in a rubidium cell and in a cæsium beam tube to 1 s in a hydrogen maser. With the values of ν_0 involved (see Table 1.1) and the linewidth originating from these processes, the atomic line quality factor, $\nu_0/\Delta\nu$, typically ranges between 10^7 and 10^9.

TABLE 1.1

Hyperfine transition frequency of Hydrogen, Rubidium 87 and Caesium

Atom	Atomic Mass	Hyperfine transition frequency (Hz)	
H	1	1 420 405 751,770	+/– 0,003
Rb	87	6 384 682 612,8	+/– 0,5
Cs	133	9 192 631 770*	

* by definition of the second.

In atomic frequency standards, the two levels of interest are determined by magnetic interaction inside the atom. It occurs between the magnetic moment of the unpaired electron of the alkali or alkali-like atoms in the ground state and the magnetic moment of the nucleus. It is called the hyperfine interaction. This interaction is weak and it leads to a small value of $E_2 - E_1$. Consequently, the resonance frequency ν_0 lies in the microwave frequency range, as shown in Table 1.1. As a practical consequence, the electronic system which controls the frequency of the associated quartz crystal oscillator can be made efficient, small and reliable.

Another annoying consequence of the smallness of $E_2 - E_1$ is that we have in fact $E_2 - E_1 << kT$, where kT is the thermal energy. Therefore, at thermal equilibrium, the two energy levels are almost equally populated. It follows that this thermal equilibrium must be broken in order to be able to observe a change in the atomic properties or an exchange of a detectable amount of energy when the atomic transition occurs. Two different methods are implemented to overcome thermal equilibrium. The first one is related to the fact that the atom shows opposite values of its effective magnetic moment whether it occupies one or the other of the two energy levels that we are considering. Consequently, the deflection of atoms in a beam passing in an inhomogeneous and intense magnetic field depends on their internal state and they can be separated [Gerlach and Stern, 1924; Gerlach, 1925]. This method of state selection is used in the cæsium beam frequency standards and in hydrogen masers. The other method is based on optical pumping [Kastler, 1950]. Besides the hyperfine interaction levels 1 and 2 in the ground state, the atoms have excited levels of much larger energy, such as level 3 shown in Fig. 1.1. An optical radiation, having an appropriate wavelength, can transfer atoms from level 2 to level 3, for instance. Spontaneous decay to levels 1 and 2 occurs in an extremely short time, < 50 ns. The net result of optical pumping is that the population of one of the ground state levels (level 1 here) increases to the detriment of that of the other level (level 2 in the example given). This is the principle of state preparation in the available rubidium frequency standard and in the optically pumped cæsium beam frequency standards presently being developed.

1.2.2.2 Passive and active atomic frequency standards

The cæsium beam and the rubidium cell frequency standards must be fed by a microwave probe signal in order to obtain the necessary information on the atomic resonance frequency. They behave as a resonator and they are called passive frequency standards. As a resonator response we obtain a discriminating resonance signal I_s which is superimposed to the total signal I_t . I_s and I_t are measures of the number of particles contributing to the signal and the background, respectively. Small size hydrogen masers with a compact cavity, but increased losses, may be operated in a similar way. The electronic system that detects the resonance frequency and controls the frequency of a 5 or 10 MHz quartz crystal oscillator is shown schematically in Figure 1.2. The probe microwave signal is synthesised from a VCXO and a frequency modulation is superimposed. This creates a modulation of the amplitude of the device response. It is processed by synchronous detection to extract the error signal which is applied to control the VCXO.

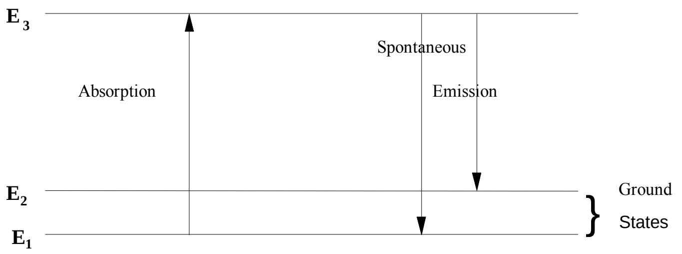


FIGURE 1.1

Principal of optical pumping

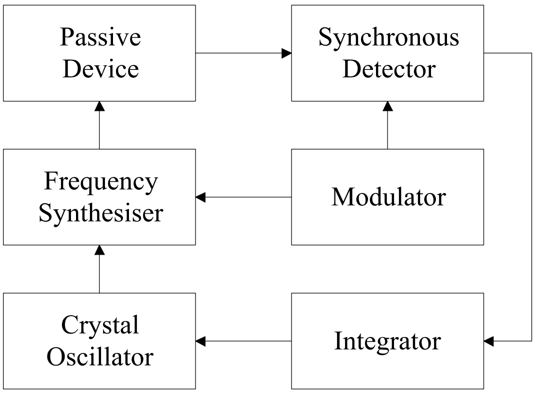


FIGURE 1.2

Block diagram representation of the frequency control of a quartz crystal oscillator in a passive frequency standard

A hydrogen maser having a full size, unloaded, cavity generates an oscillation. In this case, it is called an active frequency standard. An electronic system implements the principles of heterodyne receivers and of phase lock loops, as shown in Figure 1.3. It phase locks a quartz crystal oscillator at 5, 10 or 100 MHz to the maser oscillation. A small size hydrogen maser may be operated actively, provided that the cavity quality factor is enhanced electronically.

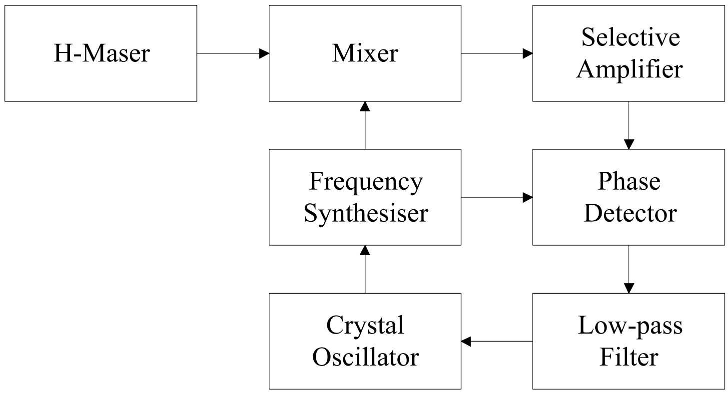


FIGURE 1.3

Block diagram representation of the phase control of a quartz crystal oscillator by the H-maser oscillation

In both cases, passive and active, the output signals of the atomic frequency standards are derived from the controlled quartz crystal oscillator.

1.3 Basic metrological concepts

1.3.1 Frequency stability

In this section, the physical origin of the random frequency fluctuations or of the frequency drifts are briefly reviewed. Typical numerical values are given in Chapter 2, Part A and Chapter 5. Measures of stability are discussed in Chapter 4. In this section, we derive theoretical values, based on the physics package of the oscillator, for the frequency stability measures developed in Chapters 3 and 4.

1.3.1.1 Definition

The term “frequency stability” is usually used instead of “frequency instability”. It is the spontaneous and/or environmentally caused frequency change within a given time interval (see Glossary). Generally one distinguishes between stochastic frequency fluctuations and systematic effects such as frequency drifts. Systematic instabilities may be caused by environmental effects. Frequency stability is dependent on the measurement system bandwidth and/or on the sample time or integration time. Random or stochastic instabilities are characterised either in the time domain or in the frequency domain or both.

1.3.1.2 Quartz crystal oscillators

1.3.1.2.1 Random frequency fluctuations

Close to the carrier, i.e. at Fourier frequencies smaller than approximately the resonator half-bandwidth, $\nu_0/2Q$, quartz crystal oscillators show flicker noise of frequency, which originates in the quartz resonator [Walls et al., 1992] and in the circuitry of the feedback loop. Then the phase noise power spectral density (PSD) varies as f^{-3}. This noise determines the frequency stability for sampling times between about 1 and 10^3s. In that range, the time domain frequency stability measure, $\sigma_y(\tau)$ (see Chapter 4), is a constant. A typical value for a 5 MHz state of the art quartz crystal oscillator is $\sigma_y(\tau) = 8 \times 10^{-14} \rightarrow 3 \times 10^{-13}$.

At Fourier frequencies larger than $\nu_0/2Q$, the oscillator is perturbed by flicker phase noise until it becomes smaller than the white phase noise. The flicker phase noise (PSD varying as f^{-1}) is a result of the processing of the flicker noise of the frequency in the oscillator feedback loop [Leeson, 1966; Sauvage, 1977].

At Fourier frequencies larger than approximately 10^2 Hz, for 5 MHz quartz crystal oscillators, the phase noise is white noise. It results from the noise added in the amplifying output circuit. The phase noise PSD within the bandwidth of the output stage is given by:

$$S_{\varphi}(f) \approx \frac{FkT}{P} \tag{1.3}$$

where k is the Boltzmann constant, T is the thermodynamic temperature, P the power generated in the oscillation loop and F is the noise factor of the output amplifier.

1.3.1.2.2 Systematic effects

For sampling times larger than $\approx 10^3$ s in the case of a 5 MHz quartz crystal oscillator, the frequency stability is perturbed by ageing effects [Walls and Gagnepain, 1992; Filler and Vig, 1993]. They yield a slow frequency change of the order of 10^{-11} per day. These ageing effects are caused by the change of the properties of the electronic components, the stress relief and the diffusion of impurities in the quartz crystal, etc. Furthermore, the frequency of quartz crystal oscillators is affected by a change of the external conditions, such as supply voltage, temperature, humidity, pressure and vibrations (see Chapter 5).

1.3.1.3 Atomic frequency standards

We shall also distinguish the short-term and medium-term frequency stability, which are determined by random processes, and the long-term stability which depends on systematic effects.

1.3.1.3.1 Short-term frequency stability

In both categories of atomic frequency standards, active and passive, the short-term frequency stability is that of the quartz crystal oscillator. This occurs for sampling times τ smaller than the control loop time constant, which is typically of the order of 1 s.

1.3.1.3.2 Medium-term frequency stability

For τ larger than the control loop time constant (≈ 1 s), up to a limit which depends on the atomic frequency standard considered (≈ 1 day for cæsium, $\approx 10^4$ s for hydrogen, $\approx 10^3$ s for rubidium), the frequency stability is determined by the quality factor of the atomic resonance and by the signal to noise ratio of its observation.

1.3.1.3.2.1 Medium-term frequency stability in passive frequency standards

The frequency stability is directly related to the precision of the measurement of the extremum of the atomic resonance pattern. It can be shown that the time-domain frequency stability measure $\sigma_y(\tau)$ is given by [Vanier and Audoin, 1989]:

$$\sigma_y(\tau) \approx \frac{1}{Q\dfrac{S}{N(\tau)}} \tag{1.4}$$

where $S/N(\tau)$ is the amplitude signal to noise ratio for the sampling time τ and Q the atomic line quality factor.

In the cæsium beam and the rubidium cell frequency standards, the noise is the shot noise of the flux of particles detected, atoms or photons, respectively. We have:

$$\frac{S}{N(\tau)} \approx \frac{I_s}{I_t^{1/2}} \tau^{1/2} \tag{1.5}$$

where I_s is the particle flux measuring the line amplitude and I_t is the total flux of particles detected. We thus obtain:

$$\sigma_y(\tau) \approx \frac{I_t^{1/2}}{QI_s} \tau^{-1/2} \tag{1.6}$$

The frequency stability varies as $\tau^{-1/2}$, which is related to the fact that the perturbing frequency noise is white noise (random and uncorrelated).

In the rubidium cell frequency standard, we usually have $I_s << I_t$. In the cæsium beam frequency standard, I_s is almost equal to I_t and equation (1.6) gives:

$$\sigma_y(\tau) \approx \frac{1}{QI_s^{1/2}} \tau^{-1/2} \tag{1.7}$$

In the passive hydrogen maser, we have:

$$\frac{S}{N(\tau)} \approx K\left(\frac{P}{kT}\right)^{1/2} \tau^{1/2} \tag{1.8}$$

where kT is the thermal noise energy in the cavity mode and P the power delivered to the cavity by the atoms. K is a constant, larger than unity, which depends on the operating parameters. We then have, in that case:

$$\sigma_y(\tau) \approx \frac{K}{Q}\left(\frac{kT}{P}\right)^{1/2} \tau^{-1/2} \tag{1.9}$$

1.3.1.3.2.2 Medium-term frequency stability in active frequency standards

In all types of oscillators, one distinguishes schematically between the frequency noise generated in the oscillation sustaining circuit and the frequency noise added to the oscillation in subsequent amplifying circuits.

The noise generated in the oscillation feedback loop is white frequency noise (in quartz crystal oscillator, this white frequency noise is usually masked by flicker noise FM). The related time-domain frequency stability measure is given by [Cutler and Searle, 1966; Vanier and Audoin, 1989]:

$$\sigma_y(\tau) \approx \frac{1}{Q}\left(\frac{kT}{2P}\right)^{1/2} \tau^{-1/2} \tag{1.10}$$

In active hydrogen masers, this white frequency noise predominates the frequency stability for τ ranging between about 20 and 10^4 s.

The noise added in the amplifying output circuit is white phase noise, as shown by equation (1.3). We have:

$$\sigma_y(\tau) \approx \frac{1}{\nu_0}\left(\frac{3FkTf_h}{2\pi P}\right)^{1/2} \tau^{-1} \tag{1.11}$$

where f_h is the noise bandwidth of the frequency stability measurement equipment. Currently, this white-phase noise is dominant for τ between approximately 1 and 10 s.

1.3.1.3.2.3 Long-term frequency stability

For τ larger than $\approx$ 1 day for cæsium, $\approx 10^4$ s for hydrogen and $\approx 10^3$ s for rubidium, the frequency stability measure $\sigma_y(\tau)$ stops decreasing when increases. It even becomes an increasing function of the sampling time τ.

There is no theory which relates the long-term frequency changes to a fundamental physical effect such as shot noise or thermal noise. Instead they are the result of slow changes of a number of frequency offsets, summarised in Chapter 2, Part A, which perturb the idealised resonance frequency defined by equation (1.1). Most of these frequency offsets depend on the way the atomic transition is observed and thus on the design or type of frequency standard. Their change is determined by ageing effects and by variations of the environmental conditions [De Marchi, 1987; Audoin et al, 1992; Mattison, 1992; Riley, 1992]. The rate of ageing and the sensitivity to external perturbations are directly related to the design choices made. Chapter 5 deals with these effects in detail.

1.3.2 Accuracy

1.3.2.1 Residual frequency offsets

Although one strives to design atomic frequency standards such that the ideal condition of atoms isolated and at rest is satisfied as much as possible, small frequency offsets cannot be avoided in current practice. The uncertainty on our knowledge of the magnitude of these frequency offsets determines the accuracy of atomic frequency standards. The physical origin of some of these frequency offsets is common to all types of frequency standards, e.g. the second order Doppler effect due to thermal motion of the atoms. Others are specific to a given type e.g. the buffer gas frequency offset and the light shift in a rubidium cell and will be dealt with in the relevant section in Chapter 2, Part A.

1.3.2.2 Definition

In general, the accuracy is the degree of conformity of a measured or calculated value to its definition (see Glossary). It is expressed as the cumulative normalised uncertainty of the value realised relative to that given by the definition. Usually, this uncertainty has two different components. In the case of atomic frequency standards, one is related to random perturbations affecting the measurement of residual frequency offsets (e.g. that due to the measurement of the magnetic field applied). The other is a systematic component that is related to a lack of knowledge of some operating parameters (e.g. the inhomogeneity of the magnetic field). These two components must be combined with care [BIPM, 1992].

1.3.2.3 Primary and secondary frequency standards

Laboratory cæsium beam frequency standards are purposely built to realise as accurately as possible the definition of the second. They actually are primary frequency standards. Their accuracy is the normalised uncertainty of the measured or estimated frequency difference between the realised value of the hyperfine transition frequency and the unperturbed transition frequency given in Table 1. The best accuracies achieved to date (1994) are about 1×10^{-14}. In the case of a frequency standard like the hydrogen maser, where the value of the unperturbed hyperfine transition frequency ν_H is known with a very small normalised uncertainty, of the order of 2×10^{12}, one may consider the concept of accuracy as well as in the case of the cæsium beam frequency standard. The hydrogen maser accuracy characterises the degree to which the output frequency can be related to the value of the unperturbed hyperfine transition frequency of the hydrogen atom. This means that the normalised uncertainty on the frequency delivered by a hydrogen maser is the combination of its accuracy and of the uncertainty on the value of the unperturbed hyperfine transition frequency of the hydrogen atom.

The accuracy of rubidium cell frequency standards is very poor because of various effects, as described below. The output frequency must be calibrated against that of a more accurate atomic frequency standard. The rubidium cell frequency standard is thus used as a secondary frequency standard.

The accuracy of quartz crystal oscillators is of the order 10^{-6} without calibration. Because of this, quartz crystal oscillators are almost always calibrated against a standard frequency, and are used as secondary frequency standards.

1.3.3 Reproducibility, resettability

The concept of reproducibility has two aspects (see Glossary):

a) with respect to a set of independent devices of the same design, it is the ability of these devices to produce the same value.

b) with respect to a single device, put into operation repeatedly without adjustments, it is the ability to produce the same value.

The resettability (see Glossary) is the ability of a device to produce the same value when specified parameters are independently adjusted under stated condition of use.

It is worth mentioning that the accuracy figure of a given atomic frequency standard (e.g. 3×10^{-12}) is an upper bound of its reproducibility and of its resettability (e.g. 5×10^{-13}) as well as of its long term frequency stability (e.g. 2×10^{-12} for the life of the device).

References

AUDOIN C. [1992] Caesium Beam Frequency Standards: Classical and Optically Pumped. *Metrologia*, Vol. 29, 113-134.

AUDOIN C., DIMARCQ N., GIORDANO V. and VIENNET J. [1992] Physical Origin of the Frequency Shifts in Cæsium Beam Frequency Standards: Related Environmental Sensitivity. *IEEE Trans. On Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 39, 412-421.

BAUCH A., DE BOER, H., FISCHER B., HEINDORFF T. and SCHRÖDER R. [1988] Long term Performance of the PTB's Primary Clock CS2. Proc. 2nd European Frequency and Time Forum, Neuchâtel, 425-435.

BESSON R.J. [1977] A New Electrodeless Resonator Design. Proc. 31st Annual Symposium on Frequency Control, Atlantic City, N.J., USA, 147-152.

BIPM [1991] "Le Système International d'Unités", Sèvres, France.

BIPM [June 1992] Guide to the Expression of Uncertainty in Measurement, Sèvres, France. ISO Document ISO/TAG 4/WG3.

BUSCA G., THOMANN P., BERNIER L.G., WILLEMIN P. and SCHWEDA H. [1990] The Classical Microwave Frequency Standards. Proc. 22nd Annual Precise Time and Time Interval (PTTI). Applications and Planning Meeting, Vienna, VA, USA, 9-17.

CLAIRON A., LAURENT P., SANTARELLI G., GHEZALI S., LEA S.N. and BAHOURA M. [1995] A Cesium Fountain Frequency Standard: Preliminary Results. *IEEE Trans. Instr. and Meas.*, Vol. 44, 128-132

CUTLER L.S. [1990] The Status of Cæsium Beam Frequency Standards. Proc. 22nd Annual Precise Time and Time Interval, (PTTI), Applications and Planning Meeting, Vienna, VA, USA, 19-27.

CUTLER L.S. and SEARLE C.L. [1966] Some Aspects of the Theory and the Measurement of Frequency Fluctuations in Frequency Standards. *Proc. IEEE*, Vol. 54, 136-154.

DE MARCHI A. [1987] New Insights into Causes and Cures of Frequency Instabilities (Drift and Long Term Noise) in Cæsium Beam Frequency Standards. Proc. 41st Annual Frequency Control Symposium, Philadelphia, USA, 53-58.

DORENWENDT K., FISCHER B. and HEINDORFF T. [1990] The PTB's Primary Time and Frequency Standards. Performance and Uncertainty. *Physica Scripta*, Vol. 41, 712-716.

ESSEN L. and PARRY J.V.L. [1957] The Caesium Resonator as a Standard of Frequency and Time. *Philosophical Transactions of the Royal Society*, Vol. 250, 45-69.

FILLER R.L. and VIG J.R. [1957] Long-Term Aging of Oscillators. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 42, 387-393.

GERBER E.A. and BALLATO A. (Editors). [1985] Precision Frequency Control, 2 volumes, Academic Press, Orlando, Florida, USA.

GERLACH W. and STERN O. [1924] Über die Richtungsquantelung im Magnetfeld. *Annalen der Physik* Vol. 74, 673-699.

GERLACH W. [1925] Über die Richtungsquantelung im Magnetfeld II. Experimentelle Untersuchungen über das Verhalten normaler Atome unter magnetischer Kraftwirkung. *Annalen der Physik*, Vol. 76, 163-197.

GIBBLE K. and CHU S. [1993] Laser-Cooled Cs Frequency Standard and a Measurement of the Frequency Shift due to Ultracold Collisions. *Phys. Rev. Lett.*, Vol. 70, 1771-1774.

HELLWIG H. [1985] Microwave Frequency and Time Standards Precision Frequency Control, Volume 2, Gerber E.A. and Ballato A. (Editors). Academic Press, Orlando, Florida, USA.

KASTLER A. [1950] Quelques suggestions concernant la production optique et la détection optique d'une inégalité de population des niveaux de quantification spatiale des atomes. Applications à l'expérience de Stern et Gerlach et à la résonance magnétique. *Journal de Physique et le Radium*, Vol. 11, 255-265.

LEE W.D., SHIRLEY J.H., LOWE J.P. and DRULLINGER R.E. [1995] The accuracy evaluation of NIST-7. *IEEE Trans. on Instr. and Meas.,* Vol. 44, 120-124.

LEESON D.B. [1966] A Simple Model of Feedback Oscillator Noise Spectrum. *Proc. IEEE*, Vol. 54, 329-330.

LEWIS L.L. [1991] An Introduction to Frequency Standards. *Proc. IEEE*, Vol. 79, 927-935.

MATTISON E.M. [1992] Physics of Systematic Frequency Variations in Hydrogen Masers. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 39, 250-255.

PRESTAGE J.D., TJOELKER R.L., DICK G.J. and MALEKI L. [1994] Progress report on the linear ion trap extended (LITE), CPEM Conference Digest, 27 June – 1 July 1994.

RILEY W.J. [1992] The Physics of Environmental Sensitivity of Rubidium Gas Cell Atomic Frequency Standards. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 39, 232-240.

SAUVAGE G. [1992] Phase Noise in Oscillators: A Mathematical Analysis of Leeson's Model. *IEEE Trans. on Instr. and Meas.*, Vol. 26, 408-410.

VANIER J. and AUDOIN C. [1989] The Quantum Physics of Atomic Frequency Standards, 2 volumes, Institute of Physics Publishing, Bristol.

VESSOT R.F.C. [1990] State of the Art and Future Direction for the Atomic Hydrogen Maser. Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Vienna, VA, USA, 29-51.

VESSOT R.F.C., MATTISON E.M., WALSWORTH R.L. and SILVERA I.F. [1990] Future Prospects for the Cold Hydrogen Maser. Proc. 4th European Forum on Time and Frequency, Neuchâtel, Switzerland, 277-281.

VIG J.R. [1991] Quartz Crystal Resonators and Oscillators for Frequency Control and Timing Applications. A Tutorial. Available from US Army Electronics Technology and Devices Laboratory. Fort Monmouth, N.J. USA.

WALLS F.L. and GAGNEPAIN J.J. [1992] Environmental Sensitivities of Quartz Crystal Oscillators. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 39, 241-249.

WALLS F.L., HANDEL P.H., BESSON R. and GAGNEPAIN J.J. [1992] A New Model of 1/f Noise in BAW Quartz Resonators. Proc. IEEE Frequency Control Symposium, Hershey, Pennsylvania, USA, 327-333.

WINELAND D.J., ITANO W.M., BERGQUIST J.C., BOLLINGER J.J., HEINZEN D.J., MANNEY C.H., MOORE F.L., RAIZEN M.G. and WEIMER C.S., [1990] Trapped-Ion Frequency Standards. Proc. 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Vienna, VA, USA, 53-60.

Chapter 2

Available frequency and time sources

Contents

Overview of Chapter 2

Whereas Chapter 1 covers the basic physics and operational concepts of precision frequency standards, Chapter 2 covers time and frequency sources. The user, considering environmental conditions and local constraints and opportunities, will be shown in this chapter different options for obtaining sources of time and frequency. The Chapter is divided into two parts. Part A deals with independent sources of time and frequency, i.e., here the available frequency standards will be described in some more detail. Part B deals with time and frequency as it may be obtained from radiating sources (both terrestrial and satellite). This, of course, includes the presentation of the relevant time and frequency transfer methods.

Typically, the sources described in Part A are frequency sources. None of them will, intrinsically, deliver UTC outside of being used as portable clocks, which is no longer the best way to obtain UTC. Part B gives sources of time (including UTC). These typically are not as good sources of frequency as those given in Part A. In this regard, it is useful to appreciate that the sources given in Part A tie to fundamental physics, from which we define frequency (the SI second). In contrast, time, which is the integral of frequency, has as its constant of integration an artefact number agreed upon by international standards bodies or as determined by the user. From such we obtain UTC, UTC(k), or some other self-consistent time scale.

Combining the best of Parts A and B in a systems approach can provide the benefits of both; i.e. a frequency source tied to the full accuracy of the SI second and an optimum estimate of UTC. This concept will be touched upon in Part B.

Part A may seem, in part, repetitious with material in Chapter 1, but the intent here is to help the user decide which of the sources given may best fill the users needs and requirements.

Part A

Local frequency and time sources

2.1 Introduction

Precision time and frequency standards based on mechanical resonances in quartz material and atomic resonance phenomena are widely used in scientific and technological applications including data networks, telecommunication systems, navigation, metrology, the world's timing systems, and scientific research such as astronomy and spectroscopy. Chapter 1 has already explained the basic principle of these devices and laid down the terminology to describe their performance. Here, each device is described in a separate section in some detail (see also [Vanier and Audoin, 1989] for a full account). Understanding the physics will better help the user see how a particular standard may perform in a given environment. Tables are added which contain relevant performance data. The most severe limitations to performance due to environmental effects are mentioned, which are treated separately in Chapter 5.

2.2 Quartz crystal frequency standards

2.2.1 The resonator

The resonator is made of a plate of quartz. Mechanical deformations of different types (flexure, extension, shear), may propagate in the bulk of the material and resonance occurs when the acoustic waves of a given type fulfil boundary conditions.

The mechanical properties of the plate can be used in an electrical circuit thanks to the piezoelectric effect, which couples electric fields and mechanical displacements in non-centrosymmetric crystals such as quartz. This material is very convenient. It can be grown in large quantities, at low cost, with sufficient chemical purity and high lattice perfection. It exhibits low electrical losses.

Figure 2.1a shows the electrical equivalent circuit of a piezoelectric resonator. In frequency and time metrology applications, the resonance frequency ν_0 is usually 5 to 10 MHz. In other applications, it can be as high as 1 GHz. The quality factor Q varies inversely with the resonance frequency. Typically, one has $Q = \frac{K}{\nu_0}$ with K ≈ 10^{13} Hz.

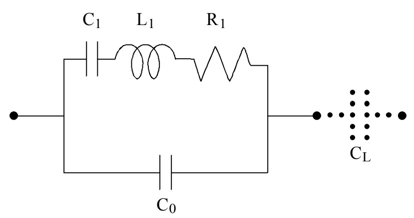


FIGURE 2.1a

Electrical equivalent circuit of a piezoelectric resonator

In general, the design of a quartz crystal resonator is aimed at minimising the perturbation of the active part of the crystal by environmental factors. For instance, the orientation of the plate with respect to the crystallographic axes, i.e. the cut, is chosen to reduce as much as possible the influence of temperature and stress upon the oscillation frequency. In an advanced design (BVA design) [Besson, 1977], the electrodes are not deposited on the faces of the vibrating plate, as is most often done. The electrodes are placed on auxiliary plates, at a distance of a few micrometers from the active one. Then, the mass loading of the main plate by the electrodes is suppressed, as well as any stress release and material migration at the interface between quartz and metal.

2.2.2 The oscillator

The resonator is inserted in a feedback loop, as shown schematically in Figure 2.1b. In steady state, the closed loop phase shift is 0 or π and the oscillation amplitude is a constant. For a given phase perturbation, the related fractional frequency change is proportional to $1/Q_L$, where Q_L is the loaded quality factor of the resonator in the network. Typically, we have $Q_L \sim 10^6$ at 5 MHz. The additional capacitor C_L makes possible the adjustment of the frequency of the oscillator. Usually, this capacitor is made of two parts added in parallel: a fixed capacitor and a varactor. The latter is used for fine tuning with an external DC voltage source. In all high quality quartz crystal frequency standards, the resonator and other temperature sensitive elements are placed in a controlled oven whose temperature is set as closely as possible at a point where the resonator frequency does not depend on temperature.

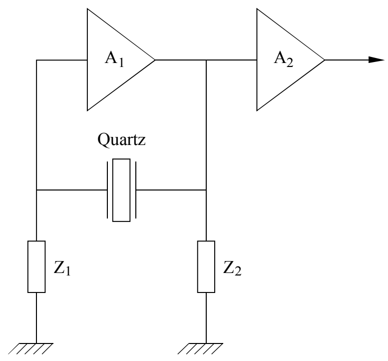


FIGURE 2.1b

Schematic representation of a quartz crystal oscillator

2.3 The rubidium gas cell frequency standard

In the rubidium gas cell frequency standard (in short: Rb standard) the state selection and detection of the clock transition is achieved using an optical-pumping method. Light from a lamp (filled with the isotope ^{87}Rb) passes a filter cell (called the hyperfine filter) which contains ^{85}Rb isotope vapour before it excites ^{87}Rb atoms in a cell filled with buffer gas inside a microwave resonator, as shown in Fig. 2.2. The buffer gas, a mixture of inert gases featuring elastic collisions with the Rb atoms, increases the interaction time, T, of the atoms with the microwave field sustained in the resonator by reducing the (inelastic) collision rate of the atoms with the walls of the cell.

The action of the isotopic filter cell alters the spectrum of the light allowing only ^{87}Rb atoms in the lower hyperfine level (E_1 in Fig. 1.1 and Fig. 2.2) to be optically pumped. This level is thus depopulated and the cell becomes transparent. If microwave radiation at $\nu_0 = 6.834$ GHz is applied to the atoms, the lower hyperfine level is again populated and optical absorption starts again. In resonance, $\nu_p = \nu_0$, the signal at the photo detector shows a minimum, as illustrated in Fig. 2.2. The spectral width of the resonance feature is typically 500 Hz due to limits on T caused by both buffer gas and residual wall collisions. The signal at ν_p is synthesised from a VCXO and modulated. Synchronous detection yields a control signal, U_R, to steer the VCXO. The centre frequency of the resonance line deviates considerably from the value of ν_0 of unperturbed ^{87}Rb atoms. Frequency shifts are caused by the presence of magnetic fields, collisions with the buffer gas, and the simultaneous interaction of atoms with optical and microwave radiation. The total shift may reach about 10^{-9} in relative units and the gas composition and light intensity is initially adjusted so as to yield optimum performance with a minimum of environmental sensitivity. The latter two effects, however, are *per se* not stable in time, as the lamp ages (change in its spectrum and its intensity) and the gas composition in the filter cell and the resonance cell itself change somewhat with time due to diffusion and outgassing. If helium is present in the environment, it can diffuse into the filter cell, and significantly change the performance of the device. Both effects also change with the ambient temperature and temperature gradients.

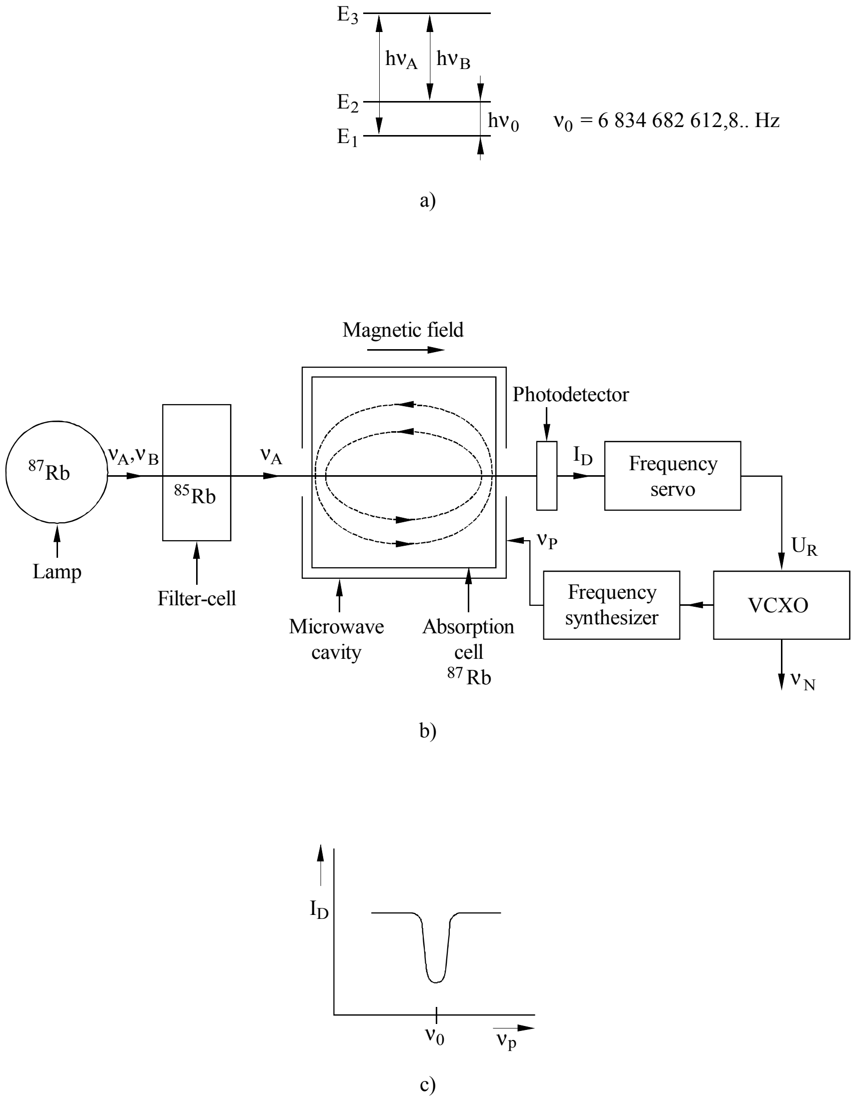


FIGURE 2.2

a) Energy levels of ^{87}Rb;

b) Rubidium standard, schematic representation;

c) Detector signal I_D versus frequency ν_p of the microwave radiation

Because the basic physics accuracy of the Rb standard is only about 10^{-9}, these standards tend to be more environmentally sensitive than cæsium and hydrogen standards. The Rb standard cannot be evaluated as a primary frequency standard (self-calibration), but instead must be calibrated against a known frequency reference. If calibrated by the manufacturer or by a standards laboratory, Rb standards can transport frequency at about the 1×10^{-11} level of accuracy. Their intrinsic short-term stability, however, can often be better than for cæsium standards because of the signal to noise ratio available in many Rb standards. Outside of vibrations and AC magnetic fields, this good short-term stability is not usually degraded by adverse environmental conditions.

Rubidium standards are available at different levels of performance, which basically depend on the size of the units and the sophistication of the control of operational parameters. In Table 2.1 small module-size units (volume less than 10^{-3} m^3) – here the hyperfine filter is often integrated into the resonance cell, containing natural Rb – and instrument size units (typical. volume of 10^{-2} m^3) are compared, and as an indication of the potentials the Rb standards used in GPS satellites which feature special temperature controls. These, however, are not commercially available.

TABLE 2.1

Performance data of rubidium gas cell frequency standards

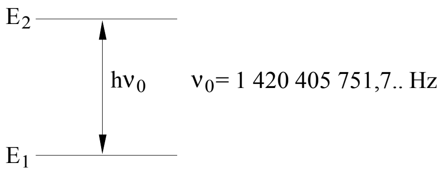
	Module (small unit)	Instrument	GPS
Accuracy	1×10^{-9}	1×10^{-10}	1×10^{-9}
Frequency instability, ($\sigma_y(\tau = 1$ s))	(2-5) $\times 10^{-11}$	(3-7) $\times 10^{-12}$	3×10^{-12}
Flicker floor ($\sigma_{y,\,min}$)	5×10^{-13}	2×10^{-13}	3×10^{-14}
Frequency fluctuations (mostly ageing, per month)	4×10^{-11}	1×10^{-11}	2×10^{-12}
Temperature (per K)	3×10^{-12}	1×10^{-12}	1×10^{-13}
Magnetic field (per 10^{-4} T)	2×10^{-11}	5×10^{-12}	2×10^{-12} (10^{-4} T = 1 Gauss)

2.4 The hydrogen maser

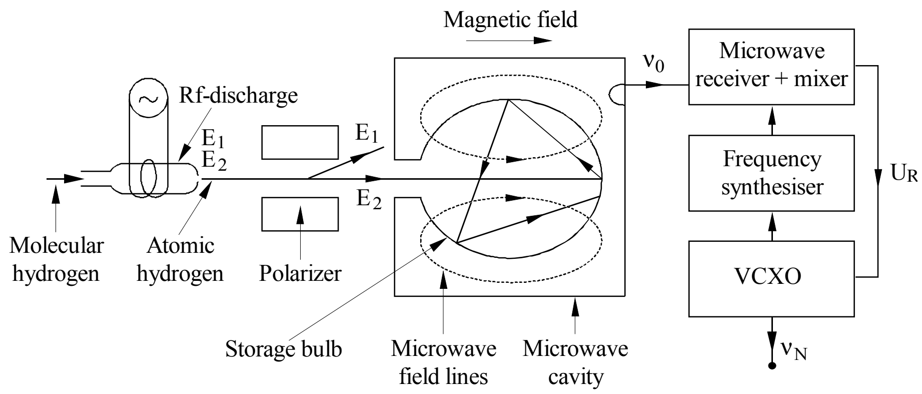
An outline of the hydrogen maser is given in Fig. 2.3. Atomic hydrogen is generated in a radio-frequency driven discharge fed by (molecular) hydrogen gas. The beam of atomic hydrogen emerges from an orifice and is magnetically state selected in a hexapole or quadrupole magnet. Hydrogen atoms in the upper energy state E_2 are directed into a storage bulb whose inner surface is coated with Teflon (Dupont trade name). The bulb is surrounded by a high-Q microwave cavity tuned to the 1 420 MHz atomic resonance frequency (see Table 1.1). While the atoms bounce around inside the bulb they decay from level E_2 to level E_1. Doing so, they release their energy through stimulated emission caused by the microwave field which the atoms themselves produce. The teflon coating ensures mostly elastic wall collisions and thus the average interaction time of the atoms with the microwave field is about one second, despite almost 10^5 collisions with the wall of the bulb [Kleppner et al., 1962; Kleppner et al., 1965]. A magnetic shield protects the interaction region from the ambient magnetic field, and a homogeneous field of about 0.1-1 μT is created in this region. Pumps are used to evacuate the device and maintain a pressure smaller than 10^{-5} Pa, despite the constant hydrogen flux.

In the ***active*** hydrogen maser the radiation emitted by the atoms themselves is used for control of the quartz frequency, but the amount of energy available from the atoms is extremely small. Hence the associated electronics for sensing this L-band signal of this quantum transition are quite elaborate. Some of the radiation, of the order 10^{-13} – 10^{-14} W, is coupled out of the cavity and fed to a low-noise microwave receiver. There the signal is mixed with a signal at a multiple frequency of the VCXO (e.g. 1 400 MHz), and the VCXO is phase-locked to the atomic radiation.

The origin of the frequency instability of the maser at short averaging times is the instability of the free-running VCXO, cavity thermal noise, and externally added noise. In spite of the complexities, with the right electronics, the excellent signal to noise available provides the best short-term stability of any frequency standard commercially available (see Table 2.2).



a)



b)

FIGURE 2.3

Hydrogen maser

a) energy levels of the hydrogen atom
b) schematic representation

TABLE 2.2

Performance data of active and passive hydrogen masers

	Active	Passive
Accuracy	10^{-12}	10^{-12}
$\sigma_y (\tau = 1$ s)	2×10^{-13}	2×10^{-12}
Flicker floor	$1 \times 10^{-14} \rightarrow 5 \times 10^{-16}$	1×10^{-14}
$\sigma_y (\tau = 1$ month) with auto-tuning without auto-tuning	 3×10^{-15} 3×10^{-13}	 1×10^{-14} Not applicable
Temperature (per K)	1×10^{-14}	1×10^{-14}
Magnetic field (per 10^{-4} T) (10^{-4} T = 1 G)	3×10^{-14}	1×10^{-14}

The long-term frequency stability is controlled largely by the frequency stability of the cavity: The cavity-pulling of the frequency is a rather severe effect in an active device, in part because of the high *Q*-value of the cavity. The active maser has therefore been traditionally considered as useful mainly for applications where its superior short- and medium-term performance is a must, for example in VLBI ground stations. Over the years, however, several methods of "auto-tuning" the cavity in the maser have been investigated and are now incorporated in some commercial devices. Some of the masers have shown a frequency ageing of as low as 10^{-16} per day compared with primary cæsium clocks [Demidov, et al., 1992; Owings et al., 1992]. The residual drift is probably caused by the ageing of the Teflon coating of the bulb, whereby the "wall-shift" factor of several 10^{-12} (caused by the collisions of the atoms with their enclosure) is changed. Because of cavity-pulling and wall-shift, the best accuracy in the determination of the unperturbed transition frequency of the hydrogen atom is about 2×10^{-12}. Though typically very small, the long-term frequency drift is still a problem in most masers.

Because of the strong impact of the mechanically defined resonance frequency of the cavity on the maser output frequency, the devices are typically sensitive to mechanical shocks and ambient temperature changes. It is only possible to achieve the long-term stability reported above if the interaction volume is surrounded by a very effective thermal shield. An active maser typically occupies a volume of 0.5 m^3, weighs 80 kg and may cost a factor of five to twenty times that of a cæsium clock.

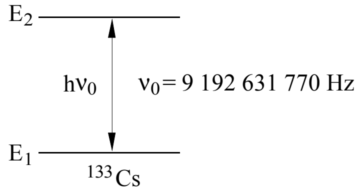
In the passive maser the cavity volume and thus the complete model can be made smaller than in an active maser. Since it is operated with a properly loaded cavity of lower Q, no self-sustained oscillation is possible. Instead, a microwave probing field is applied to the cavity and the atomic resonance is detected using a microwave receiver. The signal processing is similar to that in a cæsium clock. The short-term stability is about an order of magnitude worse than that of an active maser at $\tau = 1$ s, and the stability only improves as $\sigma_y(\tau) \sim \tau^{-1/2}$. At $\tau = 1\,000$ s, an active maser is typically more than 10 times more stable than a passive maser. However, because of the cavity servo, the passive maser is, in general, less sensitive to the environment.

2.5 The cæsium beam frequency standard

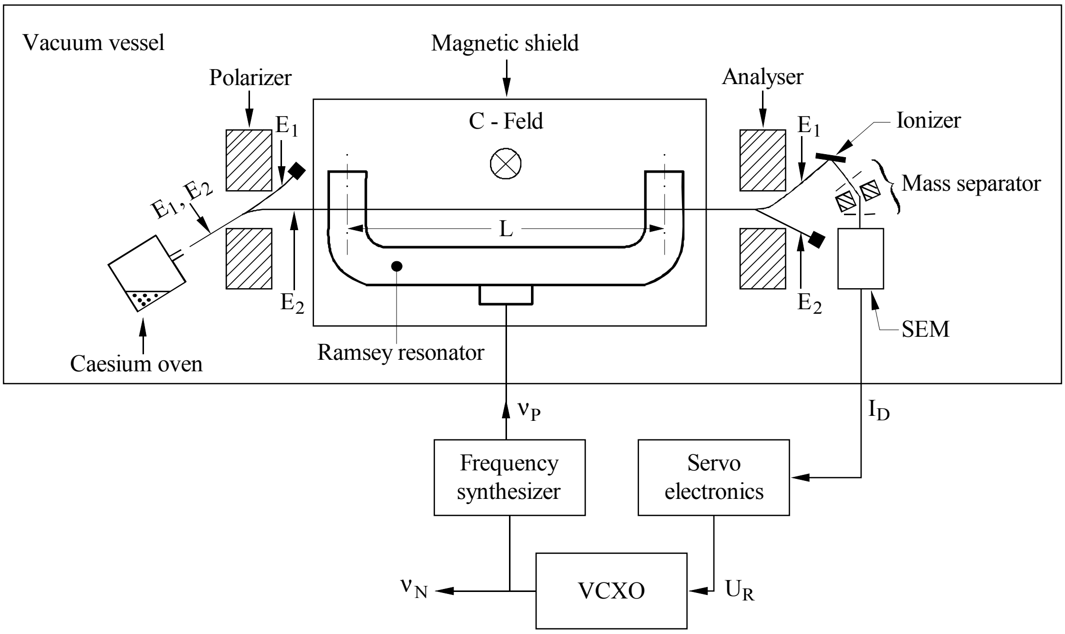
A cæsium beam frequency standard (in short cæsium clock) is shown schematically in Figure 2.4. A cæsium atomic beam effuses from an oven which contains a few grams of ^{133}Cs. The beam passes through a state-selecting magnet, named polariser, which deflects only atoms at the energy level E_2 in the desired direction. In a microwave cavity with two arms the atoms are irradiated twice with the pulse of a microwave field, following Ramsey's two oscillatory field method dating back to the fifties [Ramsey, 1950; Ramsey, 1990].

When the resonance condition (1.1) is fulfilled, the atoms are transferred to the state E_1. The analyser magnet deflects only these atoms to a hot wire detector which ionises the cæsium atoms on its surface. In commercial cæsium clocks the ions are accelerated into a mass spectrometer and directed to the first dynode of an electron multiplier. In some laboratory devices the ion current of a few picoamperes is measured directly. The output signal I_D is in both cases proportional to the number of ions being deflected by the analyser magnet. A resonance feature appears in the output signal I_D when the frequency ν_p of the probing microwave field is swept across the value ν_0. Its line width is a few hundred Hertz in commercial clocks and below 100 Hertz in laboratory standards, the value being determined by the time of flight T of the atoms through the cavity of length L (see Figure 2.4) according to (1.2). The VCXO is steered in the same way as in a Rb standard. The value of ν_0 of unperturbed cæsium atoms has been fixed as 9 192 631 770 Hz in the definition of the second.

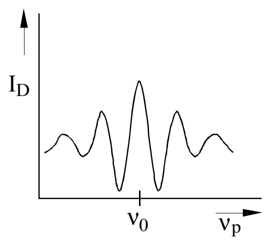
As was pointed out earlier, the concept of accuracy of a frequency standard can be applied to all three types of standards. In regard of the question, however, to what extent the SI second can be realised with a cæsium clock it seems appropriate to elaborate here the various frequency offsets to be considered. They have been considered in detail [De Marchi, 1987; De Marchi et al., 1987; Vanier and Audoin, 1989]. In brief, the most significant of them are related to:



a) Energy levels of ^{133}Cs (see Fig. 2.2)



b) Caesium clock, schematical representation



c) Detector signal I_D versus frequency ν_P of the microwave radiation

FIGURE 2.4

Cæsium clock

i) the existence of additional energy levels very close to levels 1 and 2 considered here in a simplified model. The hydrogen and the alkali atoms require an applied static magnetic field to separate a number of otherwise degenerated sub-levels. This field shifts quadratically the resonance frequency. Its value is a compromise between two opposite requirements. It must be as small as possible to limit the sensitivity of the transition frequency to its fluctuations. But it must be large enough to avoid as much as possible the overlapping of neighbouring transitions. In cæsium beam frequency standards, for instance, the value of this field is chosen around 7 μ Tesla. The related frequency offset is close to 2 Hz, i.e. 2×10^{-10} of the transition frequency and a relative variation of the magnitude of this field equal to 5×10^{-4} gives a relative change of the transition frequency of 1×10^{-13}. This leads to the implementation of: a) efficient magnetic shields in order to attenuate sufficiently the ambient magnetic field and its fluctuations, and b) a stable current source to feed the coils producing the magnetic field. The same conclusion holds for the other atomic frequency standards. The value of the static field can be measured by probing more sensitive transitions whose frequency depends linearly on its magnitude.

ii) atomic motion. It gives rise to a possible residual first order Doppler effect and to an unavoidable, relativistic, second order Doppler effect. The related normalised frequency offsets are about 1×10^{-13} in the Cs beam frequency standard, for instance.

iii) distortion of the atom wave function. Collisions between alkali or alkali-like atoms as well as collisions with a buffer gas or the walls of the container disturb the energy levels and cause a frequency shift. Similarly, collisions with the wall of a containing vessel perturb the atom hyperfine interaction and also give a frequency shift. These effects are negligible in the Cs beam frequency standard. The frequency offset is of the order of 10^{-11} in the hydrogen maser and of 10^{-9} in the rubidium cell frequency standard.

iv) state selection. A frequency shift occurs when the atoms are submitted to an extraneous radiation. This leads, for example, to the light shift, which is specific to the rubidium cell frequency standard.

v) coupling to the interrogation microwave field. As a consequence of this coupling, a sufficiently accurate tuning of the microwave cavity is required. Otherwise, a cavity frequency pulling arises. It is usually smaller than 1×10^{-13}. Proper microwave amplitude can greatly reduce the cavity pulling effects.

vi) interrogation process. In the cæsium beam and in the rubidium cell frequency standards, the measured resonance frequency is sensitive to the quality of the frequency modulated signal that is necessary to probe the transition. Furthermore, in the cæsium beam and the rubidium cell frequency standards, some of the frequency offsets depend on the amplitude of the microwave field.

Laboratory devices, referred to as primary cæsium clocks or frequency standards, have been developed in some institutes and are specifically designed to facilitate the evaluation of all systematic frequency shifts [Guinot and Azoubib, 1989]. These clocks are operated with the aim of realising the definition of the second with the highest accuracy possible. They determine the frequency of the International Atomic Time (TAI) (see Chapter 6).

The frequency instability of the cæsium clock is at short averaging times determined by the shot-noise of the atomic beam and in the long term by control of the operational parameters. The physics of Cs clocks is intrinsically less sensitive to environmental perturbations than that of Rb standards. The output of Cs standards can nevertheless be affected by adverse magnetic fields, by temperature and temperature gradients. If for example, a cæsium standard is turned upside down, its frequency will change – probably due to temperature gradients more than due to the magnetic field of the earth or that of g-forces.

The last few years have revealed new insights into the basic physical effects underlying the performance of cæsium clocks. [Bauch et al., 1988; De Marchi et al., 1984; Cutler et al., 1991; Bauch and Schröder, 1993]. These findings may have had an impact on the performance of the newest commercial models. In these devices a compromise between weight, volume, sophistication and cost of the units and their performance is unavoidable. A major break-through has occurred in the last few years in the production of a commercial Cs standard which is much less sensitive to environmental changes than heretofore. This improvement of nearly an order of magnitude in the long-term stability has come as a result of studying the physics of the transition and using its unique non-linearity coupled with digital electronics servo techniques.

Investigations into whether state selecting magnets can be dispensed with and optical pumping used instead for state selection and detection of the atoms have been in progress for some time [Audoin, 1992]. The feasibility of this method has been demonstrated and used in primary clocks [De Clercq et al., 1993; Drullinger et al., 1993] and in a bread board version of a compact cæsium clock [Petit et al., 1992], however no commercial units have been built as yet.

Commercial clocks weigh about 25 kg and fit into standard 19-inch wide cabinets (modular versions also exist). They can be simultaneously supplied with AC and DC at a power consumption of about 50 W. Primary laboratory clocks or frequency standards are not transportable or available for purchase. Table 2.3 shows performance data of commercially available standard cæsium clocks, high performance cæsium units, the new digital controlled cæsium units and the data of laboratory primary cæsium clocks as operated in national standards laboratories (constant environment) [Bauch et al., 1988; Drullinger et al., 1993].

TABLE 2.3

Performance data of cæsium clocks

Unit	Standard	High performance	Digital controlled	Primary laboratory
Accuracy	7×10^{-12}	7×10^{-12}	1×10^{-12}	1×10^{-14}
($\sigma_y(\tau = 100$ s))	3×10^{-12}	$< 10^{-12}$	$< 10^{-12}$	$< 10^{-13}$
Flicker floor ($\sigma_{y,\,min}$)	1×10^{-13}	3×10^{-14}	5×10^{-15}	3×10^{-15} *
$\sigma_y(\tau = 1$ month)	1×10^{13}	3×10^{-14}	5×10^{-15}	+
Temperature (per K)	1×10^{-13}	1×10^{-13}	$< 10^{-15}$	1×10^{-15} *
Magnetic field (per 10^{-4} T)	1×10^{-12}	1×10^{-13}	$< 10^{-14}$	+

+ does not apply to a primary clock; * estimated data (for a measurement a reference with superior stability would be needed and this is not available).

Part B
Steering references

2.6 Introduction

Scientists, engineers, and others working with time-and-frequency equipment often need to synchronise local timing signals or syntonise locally produced frequency signals with accepted national or international standards. In other cases it may be sufficient to compare the local signals with reference standards in order to determine any differences. Since the International Telecommunications Union (ITU) and other scientific and technical organisations have recommended UTC (Coordinated Universal Time) as the appropriate international reference for time-and-frequency for most applications, it is important for technical personnel working with time-and-frequency equipment to be aware of the various sources that exist for precise UTC information and the means for gaining convenient access to them.

Although the BIPM has responsibility for establishing, maintaining, co-ordinating, and generally overseeing the UTC system, users throughout the world generally access local approximations to UTC through various national time-and-frequency dissemination services that are coordinated to be within close agreement with the international UTC time scale. Recommendation ITU-R TF.685 recommends that the various local UTC(k) time scales be kept within 1 ms of UTC whenever possible. As will be described in more detail later, however, the dissemination methods used to transfer UTC from the national time-and-frequency centres to users often introduce errors that are much larger than this. Thus the level of accuracy needed at the user's site is one of the more important factors in selecting a source of time-and-frequency reference signals from among a number of alternatives.

Modern trends in time-and-frequency dissemination include the use of terrestrial and satellite-based radio broadcasts in various frequency bands, dial-up telephone links, the simultaneous exchange of time-and-frequency signals through satellite transponders, signal transfer through large-scale, synchronised digital communication networks, and links employing optical fibres, coaxial cables, or microwave systems. The various radio-based dissemination services and techniques use both dedicated frequency bands allocated by the ITU to the Standard-Frequency and Time-Signal Service (and its extension to satellite-based services) and other frequencies allocated to different radiocommunication services. As an example of the latter approach, the radiodetermination service operates various terrestrial and satellite broadcasts that must be precisely controlled with respect to time and/or frequency. Such broadcasts for radionavigation purposes are also highly useful for time-and-frequency dissemination without placing additional demands on available frequency allocations.

The ITU has allocated these specific frequencies for time-and-frequency dissemination: 20.0 ± 0.05 kHz; 2.5 ± 0.005 MHz; 5.0 ± 0.005 MHz; 10.0 ± 0.005 MHz; 15.0 ± 0.01 MHz; 20.0 ± 0.01 MHz; and 25.0 ± 0.01 MHz. The ITU Radio Regulations also include language which permits the use of specified portions of the 14-90 kHz region of the spectrum for time-and-frequency broadcasts. In addition several other frequencies were allocated to the Standard frequency and time signal-satellite service, but these have never been used for this purpose (as of 1995).

2.7 Factors to be considered in the selection and use of alternative time-and-frequency dissemination services and techniques

Beginning with the earliest radio broadcasts of standard-frequency signals in the 1920's, many different time-and-frequency dissemination services and techniques have been developed throughout the world. This resulting multiplicity of sources for standard reference signals reflects, among other things, the large variety of applications for time-and-frequency, the wide range of accuracies needed, the global nature of the coverage area to be served, differences in the relative importance of time references as compared to frequency references, and variations in the level of user sophistication and equipment costs. Today, users in many regions of the world may have access to HF broadcasts from some of the more than 20 national services using both the allocated bands and other HF frequencies; a number of low-frequency (LF) and very-low-frequency (VLF) services including both dedicated time-and-frequency broadcasts and radionavigation broadcasts such as Loran-C and Omega; several satellite broadcasts using meteorological (GOES), radionavigation (Transit, GPS, GLONASS), television, and other multi-purpose satellites (INSAT). For users needing modest timing accuracies UTC is easily available from international computer networks (e.g., INTERNET) and from convenient dial-up telephone services, particularly in Europe and North America.

In selecting one or more of the available sources for UTC time or frequency for a given application, users need to consider a variety of factors and their relative importance. These include: the availability of each service at the user's particular geographical location; whether the application requires a "time" or "frequency" reference (or both); the accuracy needed; the need for continuous availability of the reference as opposed to periodic or occasional availability; the relative importance of automatic operation; and the user costs for equipment and operation. Since no single source for UTC is optimum in all respects, compromises and trade-offs must be made when analysing these factors and the available services for a given situation. As an aid to users in selecting the best available alternative, the following sections summarise the relevant information in tabular form and then provides some further descriptive information about each of the alternatives, including brief comments on its practical use.

2.8 Comparisons of alternative sources and dissemination techniques for precise time-and-frequency references

Table 2.4 is a summary comparison of various sources of precise time-and-frequency reference information with respect to a number of factors that are important in selecting the optimum choice for any particular situation. For completeness the Table includes not only direct sources for UTC time and frequency but also a number of systems and techniques that have been found useful for the distribution and comparison of precise time-and-frequency information. The information in Table 2.4 is based on the consensus views of time-and-frequency experts participating in the work of ITU-R Study Group 7 as of mid-1993. Since some of these time-and-frequency dissemination services, systems, and techniques are evolving rapidly, potential users should maintain awareness of future developments as time goes on. It should also be noted that achieving the stated performance levels for accuracy often requires a careful calibration of the receiving equipment delays. This is particularly important at the highest stated accuracy levels.

The services, systems, and techniques summarised in Table 2.4 provide a wide range of accuracy performance from 10 ms to 1 ns for timing and from 10^{-6} to 10^{-15} for frequency. Even better results may be possible within smaller geographical regions by using optical fibre connections as indicated in the Table. As might be expected, the user costs also cover a wide range and tend to increase as the accuracy capability improves.

2.9 Additional information relating to the use of the various alternative services, systems, and techniques

Table 2.5 presents additional information about each of the alternative sources for precise time-and-frequency reference signals. This information is intended to assist potential users in the practical use of these resources. The background-information column describes the system or technique, its availability, and some of its more important advantages and disadvantages. The "Comments" column provides some brief comments on its practical use, including the type of equipment needed, methods of use, and other practical considerations.

References and Bibliography provide further information on time-and-frequency transfer techniques. The literature in this field is extensive, so the cited references are selected to be representative of what is available. The emphasis is on more recent publications which support the results quoted in the Tables, with some earlier publications which provide further background information on each particular topic.

TABLE 2.4

Characteristics of some potential sources and dissemination techniques for precise time-and-frequency reference information

Type	Typical time-transfer accuracy capability	Typical frequency transfer capability	Coverage	Availability	Ease of use	Approximate relative user cost ($US 1995)	Example system	Comments (1995)
HF broadcast	1-10 ms	10^{-6} to 10^{-8} (over 1 day)	Global	Continuous, but operator and location dependent	Depends on accuracy requirements	50 to 5,000	Many services world wide. See Rec. ITU-R TF.768	Accuracy depends on path length, time of day, receiver calibration, etc.
LF broadcast	1 ms	10^{-10} to 10^{-11}	Regional	Continuous	Automatic	3,000-5,000	See Rec. ITU-R TF.768	Depends on distance from the source and diurnal propagation (ionosphere height)
LF navigation (pulsed)	1 µs	10^{-12}	Regional	Continuous	Automatic	12,000	Loran C	Northern hemisphere coverage. Stability and accuracy based on ground wave reception.
VLF broadcast	10 ms	10^{-11} (over 1 day)	Global	Continuous	Automatic	4,000	Omega	Carrier resolution can provide better time accuracy.
Television broadcast	10 ns, common view	10^{-12} to 10^{-13} (over 1 day)	Local	Dependent upon broadcast schedule	Automatic	5,000		Calibration required for timing
Navigation satellite, broadcast	20-500 ns (See notes in Table 2)	10^{-10} to 10^{-13}	Global	Continuous	Automatic	3,000 to 15,000	GPS and GLONASS	One day averaging necessary to meet specified frequency transfer capability. Best broadcast system available today with commercial receivers.
Navigation satellite, common view	5-20 ns	10^{-13} to 10^{-15} (over 1-50 days)	Inter-continental	Continuous (calculated after the fact)	Automatic data acquisition. Requires post processing.	10,000 to 20,000 per site	GPS and GLONASS	Most accurate, widely used time synchronisation method that is available today (1995) with commercial receivers for baselines less than 8000 km.

TABLE 2.4

Characteristics of some potential sources and dissemination techniques for precise time-and-frequency reference information (continued)

Type	Typical time-transfer accuracy capability	Typical frequency transfer capability	Coverage	Availability	Ease of use	Approximate relative user cost ($US 1995)	Example system	Comments (1995)
Meteoro-logical satellite, broadcast	100 µs	Not recommended for frequency transfer	Regional (satellite footprint)	Continuous	Automatic	4,000 to 5,000	GOES	May not be available during satellite eclipse.
Geostation-ary satellite, multi-purpose broadcast	20 µs	5×10^{-10}	Regional (satellite footprint)	Continuous	Automatic	4,000	INSAT	Accuracy limited by satellite footprint. May not be available during satellite eclipse
Satellite TV	0.5-10 µs	10^{-10} to 10^{-11}	satellite footprint	Dependent on broadcast schedule	Automatic data acquisition	7,000	DBS Satellites	Without correction for satellite position
	10-100 ns	10^{-12} to 10^{-13}	Regional (satellite footprint)	Dependent on broadcast schedule	Post-processing of data required	7,000	DBS Satellites	With correction for satellite movement
Communica tion satellite, two-way	1-10 ns	10^{-14} to 10^{-15}	Regional (satellite footprint)	Continuous (as scheduled)	Data acquisition can be automatic (depending on satellite). Post processing required.	50,000 per site	North American and European networks exist.	Most precise operational method at this time.
Telephone time code	1-10 ms	10^{-8} (over 1 day)	Telephone calling range	Continuous	Automatic	100	Europe and North America	Phone line must have same path in both directions. Assumes computer and software availability.

TABLE 2.4

Characteristics of some potential sources and dissemination techniques for precise time-and-frequency reference information (continued)

Type	Typical time-transfer accuracy capability	Typical frequency transfer capability	Coverage	Availability	Ease of use	Approximate relative user cost ($US 1995)	Example system	Comments (1995)
Optical fibre	10-50 ps	10^{-16} to 10^{-17}	Local, less than 50 km	Continuous	Automatic	Transmitter and receiver $30,000 per set plus cable and underground installation costs.	Dedicated to frequency transfer	Cable must be temperature stabilised, (e.g. 1.5 m underground).
	100 ns	10^{-13} to 10^{-14} (over 1 day)	Long distance, 2000 km	Continuous	Automatic	Not applicable. The equipment is a part of a specific communication system	Synchronous Digital Hierarchy (SDH) network	Part of a digital communication system
Microwave link	1-10 ns	10^{-14} to 10^{-15}	Local	Continuous	Automatic	50,000-75,000		Sensitive to atmospheric conditions and multi-path effects. Must be 2-way to achieve stated accuracy and stability.
Coaxial cable	1-10 ns	10^{-14} to 10^{-15}	Local	Continuous	Automatic	5 to 30 per meter		Sensitive to temperature, VSWR, humidity, barometric pressure.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals

System/ technique	Background information	Comments on equipment and use
HF broadcasts	There are approximately 13 stations worldwide broadcasting on one or more of the allocated HF frequencies. Several others operate on other HF frequencies. Typical services include standard frequencies, time signals, and time intervals; time codes; voice time announcements; and UT1 time information. These services provide a convenient, easy-to-use source of UTC at modest accuracy levels. Although HF signals can be received at large distances, propagation effects can limit received accuracy and stability. Multiple stations operating on the same allocated frequencies may cause mutual interference in some areas. Reception conditions are often highly variable, depending on factors such as season, time of day, solar activity, atmospheric conditions, etc. Some HF services are being shut down in favour of other alternatives. ITU Recommendation ITU-R TF.768 contains a complete listing of HF services, including details of the content and format of the broadcasts.	Inexpensive receivers and antennas available. Diversity receivers use multiple HF frequencies to partially compensate for propagation effects. Simple short-or-long-wire antennas are often usable. Other antenna design information can be found in amateur-radio handbooks. Reception is generally better for the lower frequencies (<10 MHz) during night-time hours and for the higher frequencies (>10 MHz) during daytime hours. Reception may be intermittent due to propagation disturbances and/or interference. Optimum reception is usually during daytime or night-time hours when the ionosphere is most stable. Voice time announcements provide a few tenths of a second accuracy. For better accuracies down to about 1 ms special measurement techniques and equipment, such as oscilloscopes and electronic counters, may be required. Receiver delay calibration is also necessary for highest accuracy. Frequency-measurement accuracy is limited to about 1×10^{-7} by ionospheric motion. Beat-frequency techniques are often used along with oscilloscopes and/or counters. Frequency measurements may also be inferred from daily time-difference measurements. Calculation of signal path delays is complicated by uncertainties in the number of signal "hops" between the station and the user and the height of the reflecting layer at any point in time. Single hops can usually be assumed for distances of less than 1600 km.
LF broadcasts	This category includes broadcasts operating in the LF band (30-300 kHz) that are useful sources of UTC time or frequency but excluding navigation-system broadcasts such as Loran-C. These broadcasts are of 2 types: (1) dedicated time-and-frequency dissemination services such as DCF77, HBG, WWVB, and JJF2; and (2) stations operating in the sound-broadcasting service that have stabilised carriers and/or additional phase-or-amplitude modulations that provide coded time information. The dedicated services generally use frequencies in the 40-80 kHz range. Many of these LF broadcasts provide users with very complete time-of-year information in coded form and have found wide acceptance in many timekeeping applications. Time accuracies of less than 1 ms are	Relatively inexpensive receivers and antennas are available from commercial sources in regions served by suitable broadcasts. Commercial receivers are self-contained and provide a variety of outputs that can often be specified by the user. More sophisticated phase-tracking receivers are also available which allow users to establish direct frequency traceability to accepted sources for UTC. Typical antenna types used for these broadcasts include log-wire designs (e.g., 50-100 meters), whip antennas (e.g., 3 meters), air-loop antennas which are helpful in discriminating against interference and small ferrite-loop antennas. Reception conditions vary with the transmitter power, the user’s location, and, in some cases, the season and time of day. For longer paths between transmitter and user, avoid making measurements when there is sunrise or sunset anywhere along the path.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
LF broadcasts continued)	possible. When used as a frequency standard, LF broadcasts, including the stabilised sound broadcasts, offer calibration accuracies of less than 1×10^{-11} when averaged for about 1 day. Reliable coverage areas of the various broadcasts range from a few hundred kilometres up to 3 000 km. For more details on the available broadcasts for time-and-frequency use see Recommendation ITU-R TF.768.	Frequency calibrations of local oscillators may be performed by continuously monitoring the phase difference between the local oscillator and the received LF broadcast. Proper evaluation of the resulting phase recordings, however, requires some operator skill and experience in interpreting and accounting for various phase shifts and possible "cycle slips". Destructive interference may occur between the first-hop skywave and the groundwave, causing a sharp drop in received field intensity at certain distances from the transmitter. For a 60 kHz LF broadcast, this distance is about 1 200 km.
LF navigation broadcasts (pulsed)	Approximately 65 Loran-C stations scattered throughout the Northern Hemisphere continuously broadcast high-power navigation signals on a frequency of 100 kHz. These stations are arranged in chains of 4-5 stations each. Each chain transmits groups of precisely controlled pulses at an assigned unique Group Repetition Interval. Because the navigation signals are synchronised and syntonise by atomic standards and are carefully monitored and controlled, they can be very useful as time-and-frequency references. The Loran-C transmissions do not contain complete time-of-day information and are not a direct source of UTC time. However, if a user's clock is initially set to UTC by some other means, Loran-C can be used to keep the local clocks to within a few microseconds of UTC over long periods of time. Frequency calibrations using Loran-C can provide 1×10^{-12} accuracy when averaged over 1 day or more. Although reception from at least 3 different stations is necessary for navigation, time-and-frequency measurements require reception from only a single station.	Special Loran-C timing receivers and antennas are available commercially. The more expensive models acquire and track appropriate Loran-C signals automatically. Non-automatic receivers require significant operator experience and skill for optimum timing performance. In order to use Loran-C to keep a local clock steered to UTC an output 1 Hz pulse from the receiver can be synchronised to UTC by using "Time of Coincidence" tables published by the U.S. Naval Observatory. These tables give the specific times when the start of the Loran-C signal being received is coincident with a UTC second. Some Loran-C timing receivers can perform this synchronisation automatically. At large distances from the station the Loran-C skywave signal may sometimes be used for timing at the 50-100 ms level even when the primary groundwave signal is unusable. Seasonal effects on Loran-C propagation may cause timing variations of several microseconds. At this level receiver delays also need to be considered. Frequency calibrations of a local oscillator can be accomplished by recording the phase difference between Loran-C and the local system or by daily measurements of the phase difference using a counter. Accuracies as good as 1×10^{-12} are possible with 24-hour averaging. The development of very low cost Loran-C receivers (under $1,000) for navigation creates some possibilities for their adaptation for time-and-frequency applications, providing that the necessary technical expertise is available.
VLF broadcast	There are a number of broadcast stations operating in the 10-30 kHz range that are useful for time-and-frequency applications. These include broadcasts primarily intended for long-distance communications or navigation but which are highly stabilised in	Typical equipment used includes phase-tracking receivers, loop antennas, and chart recorders. Receiving system delays need to be calibrated for best results. Receivers used with MSK transmissions need to reconstruct a phase-coherent carrier by suitable wave multiplication and mixing. For further information on MSK signals see Note 10 to Table 2

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
VLF broadcasts (continued)	frequency and time by referencing to multiple atomic standards. Propagation is relatively stable over very large distances (thousands of kilometres), which can permit phase-tracking receivers to maintain phase to within a few microseconds over long periods of time. VLF broadcasts typically do not contain complete UTC time information and are useful primarily as a frequency reference. The Omega Navigation System is one VLF system that is useful for time-and-frequency applications. It features 8 worldwide, 10 kW transmitters providing continuous and redundant global coverage. Each station transmits the 4 navigation frequencies of 10.2, 11.05, 11.33, and 13.6 kHz sequentially in a time-shared mode. Other "unique" frequencies in the 10-13 kHz range are also transmitted by each station. Several nations also operate VLF communication stations that are useful, particularly for frequency calibration. At least some of these stations operate in an MSK (minimum shift keying) mode, requiring the use of special receiving equipment and techniques to recover a phase-stable carrier frequency.	In Recommendation ITU-R TF.768. Omega stations are located in the United States (N. Dakota and Hawaii), Japan, Argentina, France (La Reunion), Liberia, Norway, and Australia. Since each station transmits multiple frequencies in sequence, use of one of the Omega navigation frequencies for calibration requires that a commutator be used to turn the phase-tracking receiver on and off at the proper times in order to receive only the particular frequency of interest. Propagation effects often limit the useful accuracy of VLF signals, especially for very long path lengths. There are, for example, predominant diurnal and annual variations caused by ionospheric changes. Results may also be influenced by unpredictable sudden ionospheric disturbances (SID), which typically alter the ionosphere for 20-30 minutes, and by polar cap absorption (PCA) events, which alter the polar ionosphere for up to a week. In addition to the diurnal and annual variations in propagation delays at VLF, other variations been observed with periods of 27, 29.53, and 14.765 days due to various solar and lunar effects. In recent years the use of VLF broadcasts for time-and-frequency comparisons has declined due to the emergence of other systems and techniques.
Television broadcast (terrestrial links)	A number of different techniques have been tried for time-and-frequency dissemination and comparison that use television broadcast signals. These include the insertion of time-and-frequency information into the television signal, the stabilisation of television carrier frequencies and synchronisation pulses, and the common-view reception of a single television broadcast at multiple sites within a local area. The first two techniques are still in use in limited geographical areas, but the common-view reception technique is the most widely used television method. The common-view method allows the precise time comparison among multiple sites within the coverage area of a single TV single TV station. Each site simultaneously measures the time difference between a particular synchronisation pulse in the TV signal and its local clock. Subtracting the measurements from two different sites provides the difference between the local clocks plus a fixed differential propagation delay. The local clock comparisons have a typical uncertainty of about 10 ns.	Typical equipment needed includes suitable television receivers, antennas, counters, and data recorders. The television receivers must be modified to extract the particular synchronisation pulse from the received TV signal. At each measurement site arrange for the local clock pulse to start the counter and the received TV signal to stop the counter. About 10 such once-per-second measurements are usually sufficient to achieve excellent results. Since the measurements must be made simultaneously at each site and the resulting data must be exchanged, active co-operation among the sites is necessary. By making such comparisons each day over a period of time, very accurate frequency comparisons are possible based on the observed changes in the daily time differences. This assumes that the differential propagation path delay remains stable or is independently calibrated each time. The technique is especially advantageous within a limited local region because of its simplicity, relatively low cost, and high accuracy.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
Navigation satellite broadcast	There are two major satellite navigation systems in use as of 1995 which offer outstanding time-and-frequency dissemination capabilities. These are the U.S. Global Positioning System (GPS) and the Russian Global Navigation Satellite System (GLONASS). The U.S. Transit system offers a third choice, but will not be discussed further in view of its lower accuracy, higher user cost, and anticipated limited lifetime. While there are some differences between the two systems in terms of signal structure and content, use of the frequency spectrum, and satellite orbits and configuration, their similarities are much more important for time-and-frequency users. Both GPS and GLONASS employ redundant on-board atomic clocks, continuous global coverage from 21-24 operational satellites, precisely timed broadcasts which can be related to UTC (USNO) and UTC(SU), respectively, to within 100 ns, and satellite-position information included in the broadcasts which can be used for accurate path delay compensation by the user's receiver. At least 4 satellites are always in view from any location (required for navigation), but reception from only one satellite is sufficient for time-and-frequency comparison. Both GPS and GLONASS are essentially fully operational as of 1995 and provide a combined total of more than 30 satellites for time-and-frequency applications. Commercial development of receivers is proceeding rapidly with a resulting sharp decrease in user costs.	A variety of receivers are commercially available, especially for the GPS broadcasts. Some versions have been produced which can receive both GPS and GLONASS. Very small omnidirectional antennas are usually provided with the receivers. Costs have decreased sharply with the increasing demand and timing receiver packages are available in early 1995 for $3,000-$15,000. Most receivers are highly automated. During initial set-up they can be programmed to automatically track enough satellites to determine the receiver coordinates with sufficient accuracy to support sub-microsecond timing. Some care must be used in locating the antenna to minimise multipath effects. After set-up, receivers can continue to acquire and track all selected satellites in a totally automatic mode. Many receivers can be easily controlled by the user to track only certain satellites at certain times. Time differences between the received GPS signal and a local clock can often be stored in the receiver's memory for later analysis. Although the times of individual GPS and GLONASS clocks differ from the overall satellite system time which, in turn, differs from UTC, sufficient additional data are included in the satellite broadcast formats to allow a receiver to, in principle, adjust its output timing signal to be within about 100 ns of UTC(USNO) or UTC(SU). The actual display and output times, and their relationship to the relevant UTC time scales, may vary from receiver to receiver, depending on the particular manufacturer and model. Typical timing accuracies of 20-500 ns and frequency accuracies of 10^{-10} to 10^{-13} make these navigation satellite systems the best current broadcast source of highly accurate time-and-frequency for use with commercial receivers.
Navigation satellite (common-view mode)	For general background information on the GPS and GLONASS systems see the preceding entry in this Table. In the common-view mode of operation with GPS or GLONASS users at two separated sites each receives a signal from the same satellite at the same time. Subtracting the (satellite – local clock) data from the two sites provides the time difference between the local clocks. The advantage is that, in this process, variations or errors in the satellite clock are common to both paths and therefore cancel. If the SA degradation process for GPS is implemented so as to cause variations in the satellite clock, such changes do not affect the common-view measurement	Each site participating in a common-view measurement needs an appropriate GPS or GLONASS receiver and antenna, data-recording capabilities, and a communication link to other participating sites. Accurate receiver location is also required, but this can often be determined automatically by the receiver itself operating in the navigation mode. Care must be taken to ensure that the measurements extend over exactly the same time period at each site. The receiver must also be programmed to track the proper satellite that is in common view with the other sites. Typical track lengths are about 13 minutes. A subcommittee of the Consultative Committee for the Definition of the Second has recommended standard data formats and other procedural matters to facilitate the use of this method on a regular basis.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
Navigation satellite (common-view mode) (continued)	accuracy. On the other hand, if SA causes satellite-position errors to be broadcast, such errors will not be totally compensated for in the common-view measurement because each site is receiving the signal over a somewhat different path. The common-view method allows time comparison accuracies of 5-20 ns over intercontinental distances, even in the presence of SA (as it is presently implemented in early 1995). Frequency comparisons can be derived from such data to an accuracy of 10^{-13} to 10^{-15}. To facilitate common-view time comparisons among timing laboratories throughout the world the BIPM in Paris generates and distributes suitable common-view tracking schedules showing which satellites are appropriate for this method at various times.	The technique is usable for baselines between sites of up to about 8 000 km. The use of multichannel receivers in the common-view mode can provide a convenient frequency-transfer capability at the 10^{-14} level. The potential exists, for example, by using the GPS carrier phase, for extending this performance down to the 10^{-15} region by averaging over days or a few weeks. The results from many regular common-view time comparisons among national and international timing centres are published and archived by the BIPM. Receiver-system delays should be calibrated for the highest possible comparison accuracy and the antenna coordinates should be known to within less than 1 m.
Meteoro-logical satellite broadcast	Since 1974 the U.S. Geostationary Operational Environmental Satellite System (GOES) has included a time code referenced to the UTC(NIST) time scale. The time code is disseminated continuously from two geostationary satellites located normally at 75° and 135° West longitude. Satellite position data are also transmitted to users so that suitable automatic receivers can compute the signal path delay and correct their 1 Hz outputs accordingly. Specified time code accuracy as delivered to the user is 100 µs. The normal time code coverage area includes most of the Western Hemisphere with overlapping coverage of much of North and South America. Currently, as of early 1994, time code performance from the GOES/West satellite is somewhat degraded due to the use of an older, temporary satellite. Also, currently, GOES/East is operated from an orbit location of 112 degrees West longitude instead of 75 degrees. Normal satellite locations and full specified accuracy performance is expected to be restored in mid-1994 with the launch of a replacement satellite. The GOES time code includes information on the current year, day of year, hour, minute, second, UT1 correction, system accuracy, and indicators for Daylight Saving Time and leap seconds.	Commercial receivers with small antennas are available from several manufacturers. Recent versions use the transmitted satellite position information to correct for path delay and update it each 1 minute. Initial set-up requires the operator to enter the position coordinates of the receiver. The GOES time code transmissions are at 2 frequencies near 469 MHz. Because these frequencies are also allocated to the land-mobile use in the U.S., some interference, particularly near large metropolitan areas, can be expected. Receivers are reasonably effective in “flywheeling” through such periods of interference. In regions of low signal strength or frequency interference use of a simple helical or Yagi antennas may improve reception. The received time code typically shows diurnal variations with a peak-to-peak amplitude of 10-70 µs due primarily to imperfections in the software used to compute satellite position predictions. The European Meteosat system and the Japanese Geostationary Meteorological (GMS) system are basically similar to the GOES system but do not currently transmit a time code. Older GOES satellites, such as the current temporary GOES/West spacecraft, suffer time-code signal outages for about 2 hours/day during Spring and Fall eclipse periods each year. Newer satellites are not affected.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
Geostation-ary satellite, multipur-pose broadcast	The Indian INSAT geostationary satellites also transmit a UTC-referenced time code as one feature of this multipurpose system. As in the GOES case the time code signal also includes satellite-position information which allows the user to compute and compensate for the signal path delay. The INSAT satellite footprint limits primary coverage to the region of the Indian subcontinent. Within this region the timing accuracies of about 20 µs and frequency accuracies of about 5×10^{-10} are possible.	Commercial receivers are available (1993) at a cost of about $4,000. Antenna requirements are modest.
Television broadcast (satellite links)	The measurement technique is the same as reported in the case of terrestrial links, but the signals are received in common view from a direct-broadcast satellite (DBS), extending the coverage area to a nearly continental dimension. The main source of error in the determination of the clock differences arises from the variations in the position of the geostationary satellite used. This drawback can be reduced in different ways, leading to the accuracy ranges reported in Table 1. It is possible to remove the 12-hour and 24-hour periodic variations by averaging and also, most importantly, to remove the satellite longitude drift observed in the time comparisons with various techniques.	The equipment needed includes a small dish antenna, a commercial satellite TV receiver, and a TV-synchronising-pulse extractor. A time interval counter at each site measures the time differences between the local clock pulse and the received TV signal from the satellite. Two series of at least 10 such measurements, taken 12 hours apart, are needed daily. A data-acquisition system is also needed for data storage and exchange with the other stations for processing of the results. The correction for satellite-longitude drift which degrades the results can be obtained in several ways: (1) from the satellite position parameters supplied by the satellite-control station; (2) from pseudo-range measurements performed by a single station; (3) from GPS satellite measurements performed by at least 3 stations; or (4) from the time measurements performed at 3 ground stations that observe 2 satellites.
Communi-cation satellite (two-way)	At the current time the most precise and accurate method for time comparisons between remote sites is the simultaneous, two-way exchange of timing signals through communication-satellite channels. The high accuracy achievable results from the use of a two-way exchange of signals which effectively eliminates the need for precise knowledge of the satellite's position, the high degree of path reciprocity in the two directions, and the wide bandwidth of the satellite channel which permits efficient signal design.	The earth-station equipment needed at each user site must be compatible with the particular satellites being used for time transfer. Typical costs, including the necessary modems, may reach $50,000 per site. Operator skills needed for proper operation may be more stringent than for most of the other techniques discussed. Since the two-way technique is essentially a point-to-point communication system, it should not be regarded as a general dissemination technique.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
Communi-cation satellite (two-way) (continued	One disadvantage of the technique is the need for each site to both transmit and receive signals and then to exchange the data for post-processing. The earth-station equipment at each site tends to be rather expensive, especially if the system is highly automated. Participants in the time transfers must coordinate with each other and with the satellite system operator. Because of the potential accuracy of near 1 ns and the precision of 0.1-0.5 ns, many timing laboratories in various parts of the world are developing a two-way time-transfer capability. Special modems are being developed which are optimised for high accuracy and long-term stability. Suitable satellite channels appear to be available throughout the world at reasonable cost.	As typically implemented, two or more sites exchange timing signals on a regular basis, such as once or twice per week. Because of the inherent accuracy and stability of the method, it is usually only necessary to perform the exchanges for a few minutes per time. The measurement process involves measuring the difference between the satellite-signal arrival time and the local clock. Such measurements are often made once per second for a period of a few minutes. Subtraction of the simultaneous measurements at each site, divided by 2, provides the difference between the site clocks. For the highest achievable accuracy of 1-10 ns it is important to calibrate the signal delay through the ground-station equipment. This may be a difficult problem since the relevant quantity needed is the difference between the delays through the transmit and receive portions of the system. Several specialised techniques have been developed for this purpose. Depending on the particular satellite system being used and the locations of the e stations, extensive administrative procedures may be required in order to certify the earth station equipment and gain acceptance for satellite access.
Telephone time code (two-way)	A number of timing centres in Europe and N. America have established services designed to disseminate coded UTC time information over telephone lines in an automated mode. Typically, computers or other automated systems are programmed by the user to dial such services as needed, receive an ASCII time code from the timing centre, reset the local clock to the correct time, and, in some cases, to automatically compensate for the path delay through the telephone link. Depending on the particular service, the path delay compensation can be performed either by the timing centre's equipment or at the user's site. The compensation for delay is based on measurements of the round-trip delay time and assumes that the path is reciprocal. Time-transfer accuracies of 1-10 ms are possible, even when satellite links may be involved. In addition to the UTC time of day, most services established to date also include information on the year, day of year, UT1 corrections, leap second warnings, and indicators for daylight-saving time.	Equipment requirements to use such services are minimal. Aside from the computer or other equipment containing the clock to be set, only a suitable modem, access to a telephone line, and clock-setting software is needed. In order to perform the path delay compensation the user may also need to be able to echo the received signal back to the timing centre. Usually, a telephone connection time of only a fraction of a minute is needed to perform a satisfactory time transfer. Software for using such services is relatively simple to develop by users or some versions of example software are often available via computer bulletin boards, from the timing centres, or from commercial sources at reasonable cost. Most of the available telephone services can also be used on a one-way mode where there is either no compensation for path delay or a fixed, average delay is used. Accuracy for this mode may be in the range of 0.1 to 0.5 s. By making periodic measurements of a local clock using one of the telephone services, an average frequency can be determined. Accuracies of about 10^{-8} are possible with 1-day averages.

TABLE 2.5

Additional information relating to the practical use of the various alternative sources of time-and-frequency signals (continued)

System/ technique	Background information	Comments on equipment and use
Optical fibre	Optical fibres offer excellent potential for transferring time-and-frequency signals with very high accuracy over both short (<50 km) and long distances. While dedicated UTC dissemination services using optical-fibre distribution do not currently exist, the technique is included here in recognition of its future potential. Two types of fibres, multimode and single mode, are in use today. Multimode fibre is generally used to transmit digital data and low frequencies over a relatively short distance (e.g., 1 km). Single-mode fibre is best for longer distances (e.g., 50 km) and supports wide bandwidth (e.g., 5 MHz to 100 GHz). Single-mode fibre with a 1 300 nm laser is required to meet the performance given in Table 1 for local distances. The accuracies stated in Table 1 for long fibre-optic links has been achieved in a digital telecommunications system adhering to CCITT Recommendations G.707, 708, and 709 over a distance of 2 400 km. This particular system was designed to meet ITU-T requirements as well as to perform time-and-frequency-transfer experiments.	In a practical implementation of a fibre-optic link for time-and-frequency transfer at highest possible accuracy levels it is important to stabilise the temperature of the cable. The nominal coefficient of delay with respect to temperature is 7 PPM/^{o}C. In order to meet the stated performance in Table 1 for links longer than 50 km the cable should be put underground to a depth of at least 1.5 m. For a dedicated optical-fibre link for time-and-frequency transfer, the cost is about $30 000 per site for transmitters and receivers plus the cost of the cable and its underground installation. Insertion loss is 0.5 dB/km. Potential users and suppliers of UTC should maintain current awareness of the development of regional, national, and international digital synchronised telecommunication networks. Such networks may provide an excellent, convenient means for distributing high-accuracy UTC time-and-frequency in the future.
Microwave link	The use of microwave links to distribute time-and-frequency within local areas can provide accuracies as high as 1-10 ns for timing and 10^{-14} to 10^{-15} for frequency when used in a two-way mode.	Equipment is relatively expensive ($50 000 – $75 000). Results are sensitive to atmospheric conditions and multipath effects. For highest accuracy two-way operation is required with a continuously operating feedback loop for nulling out phase variations.
Coaxial cable	Coaxial cables offer a convenient means of transferring time-and-frequency information over distances of less than several hundred meters. To achieve the accuracy performance given in Table 1 careful attention must be paid to temperature environment, temperature stability, and the type and length of cable. Good temperature stability can be achieved by burying the cable at least 1.5 m underground.	Cable cost is about $5 – $30 per metre. Insertion loss is dependent on cable length, type, and the frequency used. Solid-dielectric cable has a coefficient of delay of 250 PPM (or even greater at 25 ^{o}C). Air dielectric is 15 PPM, but must be dry-nitrogen pressurised with a dual-stage pressure regulator in an environment controlled to within 1^o C.

References

AUDOIN C. [1992] Caesium Beam Frequency Standards: Classical and Optically Pumped. *Metrologia*, Vol. 29, pp. 113-134.

BAUCH A., DE BOER H., FISCHER B. HEINDORFF T., SCHRÖDER R. [1988] Performance of the PTB's Primary Clocks CS2 and CS1. Proc. 42nd Annual Frequency Control Symposium. IEEE Cat. CH2588-2/88, pp. 490-495.

BAUCH A. and SCHRÖDER R. [1993] Frequency Shifts in a Cesium Atomic Clock due to Majorana Transitions. *Annalen der Physik 2*, pp. 421-449.

BESSON R.J. [1977] A New Electrodeless Resonator Design. Proc. 31st Ann.Symp.Frequ.Control, pp. 147-152.

CUTLER L.S., FLORY C.A., GIFFARD R.P., DE MARCHI A. [1991] Frequency Pulling by Hyperfine σ Transitions in Cesium Beam Atomic Frequency Standards. *J.Appl. Phys*. Vol. 69, pp. 2780-2792.

DE CLERCQ E., ROVERA G.D., BOUZID S., CLAIRON A. [1993] The LPTF Optically Pumped Primary Frequency Standard. *IEEE Trans. Instr. Meas*, IM-42, pp. 457-461.

DE MARCHI A.,[1987] New Insights into Causes and Cures of Frequency Instability (Drift and Long Term Noise) in Cesium Beam Frequency Standards. Proc. 41st Annual Frequency Control Symposium, Philadelphia, USA, pp. 53-58.

DE MARCHI A., ROVERA G.D., PREMOLI A. [1984] Pulling by Neighboring Transitions and its Effects on the Performance of Caesium-Beam Frequency Standards. *Metrologia*, Vol. 20, 37-47.

DE MARCHI A., ROVERA G.D. and PREMOLI A. [1987] Effects of Servo-Loop Modulation in Atomic Beam Frequency Standards Employing a Ramsey Cavity. *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 34, 582-591.

DEMIDOV N.A., EZHOV E.M., SAKHAROV B.A., ULJANOV B.A., BAUCH A., FISCHER B. [1992] Investigations of the Frequency Instability of the CH1-75 Hydrogen Maser. Proc. 6th European Frequency and Time Forum, esa SP-340, pp. 409-414.

DRULLINGER R.E., SHIRLEY J.H., LOWE J.P., GLAZE D.J. [1993] Error Analysis of the NIST Optically Pumped Primary Frequency Standard. *IEEE Trans. Instr. Meas*. IM-42, pp. 453-456.

GUINOT, B. and AZOUBIB J. [1989] Comparison of Primary Frequency Standards, De Marchi A. (Ed.). Frequency Standards and Metrology, Berlin, Heidelberg, Springer-Verlag, pp. 37-43.

KLEPPNER D., BERG H.C., CRAMPTON S.B., RAMSEY N.F., VESSOT R.F.C., PETERS H.E. and VANIER J. [1965] Hydrogen-Maser Principles and Techniques. *Physical Review*, Vol. 138, pp. A-972-A983.

KLEPPNER, D., GOLDENBERG H.M. and RAMSEY N.F. [1962] Theory of the Hydrogen Maser. *Physical Review*, Vol. 126, pp. 603-615.

OWINGS H.B., KOPPANG P.A., MACMILLAN C.C., PETERS H.E. [1992] Experimental Frequency and Phase Stability of the Hydrogen Maser Standard Output as Affected by Cavity Auto-Tuning,. *Proc. IEEE Frequency Control Symposium.* IEEE Cat.No. 92CH3083-3, pp. 92-103.

PETIT P., GIORDANO V., DIMARCQ N., CEREZ P., AUDOIN C., THEOBALD G. [1992] Miniature Optically Pumped Cesium Beam Resonator. Proc. 6th European Frequency and Time Forum, ESA SP-340, pp. 83-86.

RAMSEY N.F. [1950] Molecular Beam Resonance Method with Separated Oscillating Fields. *Physical Review*, Vol. 78, 695-699.

RAMSEY N.F. [1990] Experiments with separated oscillatory fields and hydrogen masers. *Reviews of Modern Physics*, Vol. 62, 541-552.

VANIER J. and AUDOIN C. [1989] The Quantum Physics of Atomic Frequency Standards. Bristol and Philadelphia, Adam Hilger.

Bibliography

ALLAN D. W. [1990] Remote time and frequency comparisons now and in the future. Proc. 4th European Frequency and Time Forum, p. 619.

BEEHLER R. [1982] GOES satellite time code dissemination. Proc. 14th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Washington, DC, USA, December, pp. 57-87.

CORDARA F., PETTITI V., QUASSO R. and RUBIOLA E. [1992] Performances of a date dissemination code on telephone lines using commercial modems. Proc. 24th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, McLean, VA, USA, December, pp. 243-253.

CUBBAGE R. W. [1992] SONET synchronization: what's happening. Proc. 24th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, McLean, VA, USA, December, pp. 337-343.

DALY P., KOSHELYAEVSKY N. B., LEWANDOWSKI W., PETIT G. and THOMAS C. [1993] Comparison of GLONASS and GPS time transfers. *Metrologia*, Vol. 30, 2, 89-94.

DAVIS J. A., PEARCE P. R., KIRCHNER D., RESSLER H., HETZEL P., SOERING A., DE JONG G., GRUDLER P., BAUMONT F. and VEENSTRA L. [1994] Two-way satellite time transfer experiments between six European laboratories using the INTELSAT (VA-F13) satellite. Proc. 8th European Frequency and Time Forum, p. 297.

DAVIS J. A., PEARCE P. R., MENARY P., HUTCHINS M. and SAGIN M. [1991] Time transfer by passive use of direct broadcasting satellite signals. Proc. 5th European Time and Frequency Forum, p. 74.

DE JONG G. [1992] Prospects of two-way international time scale comparisons. Proc. 6th European Frequency and Time Forum, p. 69.

GAEDE H. J. [1988] Precise determination of UTC in remote regions using OMEGA signals with reference to monitoring performance of a hydrogen maser. Proc. 2nd European Frequency and Time Forum, p. 365.

GUINOT B., LEWANDOWSKI W. and THOMAS C. [1990] A review of recent advances in GPS time comparisons. Proc. 4th European Frequency and Time Forum, p. 307.

HETZEL P. [1988] Time dissemination via the LF transmitter DCF 77 using a pseudo-random phase-shift keying of the carrier. Proc. 2nd European Frequency and Time Forum,. p. 351.

ITU-R Recommendations [1994] TF Series Volume, Time signals and frequency standards emissions**.**

ITU-R Recommendations [1995] TF Series Fasicle, Time signals and frequency standards emissions**.**

JESPERSEN J. L., BLAIR B. E. and GATTERER L. E. [1972] Characterization and concepts of time-frequency dissemination. *Proc. IEEE*, Vol. 60, 5, 502-521.

KALLIOMAKI K. and MANSTEN T. [1993] TV-frequency standard as a traceability source in accredited calibration laboratories. Proc. 7th European Frequency and Time Forum, p. 383.

KIHARA M. and IMAOKA A. [1992] Timing and time signal distribution in digital communication networks. Proc. 6th European Frequency and Time Forum, p. 489.

KIRCHNER D. [July, 1991] Two-way time transfer via communication satellites. *Proc. IEEE*, Vol. 79, 7, 983-991.

KIRCHNER D., RESSLER H., GRUDLER P., BAUMONT F., VEILLET CH., LEWANDOWSKI W., HANSON W., KLEPCZYNSKI W. and ULRICH P. [1993] Comparison of GPS common-view and two-way satellite time transfer over a baseline of 800 km. *Metrologia*, Vol. 30, 183-192.

KITCHING I. D. and DALY P. [1989] Time references from GLONASS satellites. Proc. 3rd European Frequency and Time Forum, p. 121.

LEVINE J., WEISS M., DAVIS D. D., ALLAN D. W. and SULLIVAN D. B. [1989] The NIST Automated Computer Time Service. J. Res. NIST, Vol. 94, 5, 311-321.

LEWANDOWSKI W. and THOMAS C. [1991] GPS time transfer. *Proc. IEEE*, Vol. 79, 7, 991-1001.

MEYER M., GRANVEAUD M., LAPORTE B., VERNOTTE F. and VINCENT M. [1993] Improved time transfer using geostationary direct TV satellites. Proc. 7th European Time and Frequency Forum, p. 129.

NIST [1990] Time and frequency users manual. National Institute of Standards and Technology Special Publication 559, available from the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402 (Note: This publication contains information at a tutorial level on how to use many of the techniques and services mentioned in this chapter).

POTTS C. E. and WIEDER B. [1972] Precise time and frequency dissemination via the LORAN-C system. *Proc. IEEE*, Vol. 60, 5, 530-539.

PUSHKIN S. B. [1993] GLONASS common-view time transfer. Proc. 7th European Frequency and Time Forum, p. 147.

SEN GUPTA A., HANJURA A. K. and MATHUR B. S. [1991] Satellite broadcasting of time and frequency signals. *Proc. IEEE*, Vol. 79, 7, 973-983.

SENNEDOT D., MINGUY Y. and THOMAS B. [1993] Use of Loran-C signals for time/frequency transfer. Proc. 7th European Frequency and Time Forum, p. 423.

STARKER S. [1994] Time transfer via satellites – a survey. Proc. 8th European Frequency and Time Forum, p. 275.

STEBBINS S., GIFFORD A. and BEARD R. L. [1989] A comparison of time transfer techniques. Proc. 21st Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Redondo Beach, CA, USA, November, pp. 215-222.

SYDNOR R. L. and CALHOUN M. [1993] Precise frequency distribution using fiber optics. Proc. 7th European Frequency and Time Forum, p. 399.

Chapter 3

Characterisation: frequency domain, time domain

Contents

3.1 Introduction

This chapter is an introduction to the methods used to characterise the random phase and frequency instabilities of oscillators in both the frequency domain and the time domain.

The characterisation of phase and frequency instabilities in precision oscillators is already covered in Recommendation ITU-R TF.538 and extensively in the literature. For example references [Allan, 1987; Allan, et al., 1988; IEEE, 1988; Lesage et al., 1979; Rutman, 1978 and Stein, 1985] are all review papers covering the similar topics to the present Chapter.

The originality of this particular presentation is the fact that the different topics are presented in a unified approach using electrical engineering concepts such as the analytic signal, random processes, linear operators, transfer functions, the sampling theorem etc. instead of following the conventional approach inherited from the historical development of time and frequency characterisation concepts.

A particular effort has been made in order to give to the reader the necessary insight into the theoretical aspects of the phase and frequency noise characterisation methods and at the same time to give the practical tips and algorithms necessary for their application. The next chapter covers similar topics to the above, but with more experimental guidance.

3.2 Model of the oscillator

3.2.1 The phasor model and the analytic signal

The signal of the sinusoidal oscillator is a narrow-band, band-pass signal, centred about the carrier frequency ν_0. It can be modelled by a phasor, i.e. by a vector rotating in the complex plane at the carrier frequency:

$$Y(t) = A\gamma\,(t)\ \exp\,(\mathrm{j}2\pi\nu_0\, t) \tag{3.1}$$

A is the nominal amplitude, $\gamma(t)$ is the complex envelope and exp $(\mathrm{j}2\pi\nu_0\, t)$ the phasor. The real signal $s(t)$ is the real part of the phasor signal $\psi(t)$:

$$Y(t) = s(t) + \mathrm{j}\bar{s}(t) \qquad \text{where } s(t) \text{ is the Hilbert transform of } \bar{s}(t) \tag{3.2}$$

i.e. the projection of the phasor on the real axis.

It can be shown that the phasor representation of a bandpass signal is a particular case of an analytic signal [Bedrosian, 1962; Bernier et al., 1985].

The complex envelope $\gamma(t)$ of the analytic signal is a low-pass complex random process that represents the information modulated on the carrier. In the case of a telecommunication band-pass signal the low-pass message signal is translated on purpose about a carrier frequency by modulation. In the case of an oscillator signal, on the other hand, the complex envelope is a random process characteristic of a certain type of oscillator: quartz oscillator, atomic frequency standard etc. A review of random modulation can be found in [Papoulis, 1983].

The complex envelope can be separated either into an in-phase $p(t)$ and a quadrature $q(t)$ low-pass processes or into a unity carrier, an amplitude low-pass noise process $\varepsilon(t)$ and a phase low-pass noise process $\phi(t)$:

$$\gamma(t) = p(t) + \mathrm{j}q(t) = [1 + \varepsilon(t)]\ \ \exp\,(\mathrm{j}\phi(t)) \tag{3.3}$$

3.2.2 The low-noise oscillator

In the case of a precision or low-noise oscillator the mean square value of the phase noise process is much smaller than unity, and the mean square of the amplitude noise process is much smaller than unity,

$$E\{\phi^2(t)\} << 1$$

$$E\{\varepsilon^2(t)\} << 1$$

and where $E\{\}$ is the statistical expectation operator and the phase noise exponential can be linearised:

$$\exp(\mathrm{j}\phi(t)) \approx 1 + \mathrm{j}\,\phi(t)$$

yielding to lowest order:

$$\Psi(t) = A\,[1 + \varepsilon(t) + \mathrm{j}\,\phi(t)\,]\exp(\mathrm{j}2\pi\nu_0 t) \tag{3.4}$$

Under these conditions the in-phase component, $p(t)$, can be identified as a unity carrier plus the amplitude noise while the quadrature component, $q(t)$, is identified as the phase noise.

$$\begin{aligned} p(t) &= 1 + \varepsilon(t) \\ q(t) &= \phi(t) \end{aligned} \tag{3.5}$$

The phase noise and the amplitude noise are orthogonal when the complex envelope is represented in the complex plane. This property enables the independent demodulation of the phase and amplitude noise processes by synchronous demodulation as explained below, but the property holds true only in the low-noise case.

3.2.3 Spectrum of the low-noise oscillator

Note that the definitions relevant to the autocorrelation functions and to the power spectral densities are given in the appendix on random processes and follow the notations used in [Wozencraft et al., 1965]. In particular $S_{\phi\phi}(f)$ refers to the two-sided power spectral density of the random process $\phi(t)$, $S^{+}_{\phi\phi}(f)$ refers to the one-sided power spectral density of the random process $\phi(t)$, while $R_{\phi}(\tau)$ refers to the autocorrelation function of the random process $\phi(t)$. Note, moreover, that outside the present chapter of the handbook, the autocorrelation function of $\phi(t)$ is noted $R_{\phi}(\tau)$ and the corresponding spectral density $S_{\phi}(f)$ The ITU and IEEE standard is to use the one-sided spectral densities. The autocorrelation function of the complex envelope associated with the low-noise oscillator signal as defined by (3.5) in [Bernier et al., 1985]:

$$R_{\gamma\gamma}(\tau) = 1 + R_{\varepsilon\varepsilon}(\tau) + R_{\phi\phi}(\tau) \tag{3.6}$$

The power spectral density of the oscillator signal s(t) is the Fourier transform of the low-pass autocorrelation function (3.6) translated about the carrier frequency, because of (3.1) and (3.2), yielding:

$$S^{+}_{ss}(f) = \frac{A^2}{2}\Big(\delta(f - \nu_0) + S_{\varepsilon\varepsilon}(f - \nu_0) + S_{\phi\phi}(f - \nu_0)\Big) \tag{3.7}$$

where $\delta(f - \nu_0)$ is a Dirac function representing the carrier and where $S_{\varepsilon\varepsilon}(f - \nu_0)$ and $S_{\phi\phi}(f - \nu_0)$ are the two-sided power spectral densities of the amplitude and phase noise processes.

The low-noise oscillator therefore appears on the spectrum analyser as a carrier at the nominal frequency ν_0 superposed to the amplitude and phase two-sided power spectral densities translated about the carrier frequency. In telecommunication terms this spectrum is that of a carrier AM modulated and simultaneously PM modulated with a low modulation index by the amplitude and phase random processes characteristics of the oscillator.

The equations in this section are true only if the signal spectrum is symmetrical about ν_0. If not, then (3.6) has a cross term, and (3.7) can be written as (3.7′){[Bernier 1985, equation 14}. For $|\nu| < \nu_0$;

$$\tfrac{1}{2}\Big[S^{+}_{ss}(\nu_0 + f) + S^{+}_{ss}(\nu_0 - f)\Big] = \tfrac{A^2}{4}\Big[S_{\gamma\gamma}(f) + S_{\gamma\gamma}(-f)\Big] = \tfrac{A^2}{2}\Big[\delta(f) + S_{\varepsilon\varepsilon}(f) + S_{\varphi\varphi}(f)\Big] \tag{3.7′}$$

which says that the average sideband level=AM level + PM level. Note that $S_{\varphi\varphi}(f) = \mathcal{L}(f)$, see (3.24).

3.2.4 The high-noise oscillator

In the case of a high phase noise oscillator the mean quadratic value of the phase noise process is large:

$$E\{\phi^2(t)\} \geq 1$$

and the exponential term of the complex envelope cannot be linearised. In this case the phase noise process contributes to both the in-phase and quadrature components of the complex envelope yielding:

$$\begin{aligned} p(t) &= (1+\varepsilon(t)) \cos\phi(t) \\ q(t) &= \sin\phi(t) \end{aligned} \tag{3.8}$$

In this case the amplitude and phase noise processes are not orthogonal anymore in the complex plane. The lack of orthogonality complicates the problem of independent demodulation of the processes. The demodulation problem is treated in the Chapter 4.

3.2.5 Spectrum of the high-noise oscillator

Assuming that the amplitude noise is negligible and that the phase noise is high, it can be shown [Bernier et al., 1985] that the autocorrelation function of the complex envelope can be separated into three terms corresponding respectively to the carrier and to the in-phase and quadrature components [Bernier et al., 1985]:

$$R_{\gamma\gamma}(\tau) = \exp(-R_{\phi\phi}(0)) + r_p(\tau) + r_q(\tau) \tag{3.9}$$

where:

$$\begin{aligned} r_p(\tau) &= [\exp(-R_{\phi\phi}(0)[\cosh(R_{\phi\phi}(\tau)) - 1] \\ r_q(\tau) &= [\exp(R_{\phi\phi}(0))]\sinh(R_{\phi\phi}(\tau)) \end{aligned} \tag{3.10}$$

The one-sided power spectral density of the band-pass signal $s(t)$ is therefore [Bernier et al., 1985]:

$$S_{ss}^{+}(f) = \frac{A^2}{2}\left(\exp(-R_{\phi\phi}(0))\,\delta(f - \nu_0 + G_p(f-\nu_0) + G_q(f-\nu_0)\right) \tag{3.11}$$

where G_p and G_q are the Fourier transforms of r_p and r_q.

As the mean quadratic value $R_{\phi\phi}(0) = E\{\phi^2(t)\}$ of the phase noise increases beyond unity, the carrier vanishes exponentially, i.e. the signal becomes incoherent, and the in-phase and quadrature spectral components both tend towards a Gaussian shape regardless of the original shape of the phase noise power spectral density. This is known in telecommunication theory as Woodward's theorem [Bernier et al., 1985; Blachman et al., 1969; Papoulis, 1983]. In the case of a low-pass white phase noise process, the carrier disappears completely and the spectrum becomes Gaussian when the mean squared value of the phase noise exceeds about 8 rad^2 [Bernier et al., 1985]. The above is a simple stationary case. However, in any practical oscillator the phase noise is indeed non-stationary. The non-stationary case is treated in [Walls et al., 1975].

3.2.6 Effect of frequency multiplication

When a carrier frequency is multiplied by n in a frequency multiplier, the amplitude noise $\varepsilon(t)$ can be kept very low by the use of limiters but the phase noise $\phi(t)$ is multiplied by n. The mean squared value of the phase noise process is therefore multiplied by n^2 in the ideal case. In practice, the phase noise level may be higher than this fundamental lower limit and the amplitude noise level depends on the actual characteristics of the multiplier device. If the multiplication factor, n, is high enough, the mean squared value of the phase noise process will eventually reach and exceed unity. If this is the case the original low-noise oscillator signal becomes a high-noise oscillator signal, i.e., becomes more or less incoherent as described above. See Chapter 4 for more details and for the demodulation of the noise process.

3.2.7 Demodulation of the noise processes

The application of both frequency and time domain characterization methods described in the next sections implies that the low-pass phase noise process $\phi(t)$ of the oscillator under test can be extracted by demodulation. Extraction of the phase noise is simpler in the case of a low phase noise oscillator.

For the characterization of oscillators in the time domain, the demodulation method is heterodyne down-conversion to a relatively low IF frequency followed by zero-crossing detection. The zero-crossings are then "time tagged " or the IF period is averaged and sampled using a digital period/frequency counter or a time-tag counter.

In the case of spectral domain characterization methods the demodulation technique is synchronous detection. Synchronous detection of a band-pass signal with the local oscillator in-phase with the carrier yields the in-phase component of the complex envelope i.e. $p(t)$. Conversely the synchronous detection of a band-pass signal with the local oscillator in quadrature with the carrier yields the quadrature component of the complex envelope i.e. $q(t)$.

As shown above the amplitude noise and the phase noise processes are orthogonal in the complex plane, and thus can be easily separated by synchronous detection in the case of a phase noise process with a small mean squared value.

Moreover at low Fourier frequencies the phase noise of any real oscillator is usually modeled by a non-stationary process of the "power-law" type, see sections about the polynomial model and about non-stationary processes below, whose mean value can be defined only locally. Therefore, the only way to maintain the local oscillator in quadrature with the carrier for the purpose of the synchronous detection of the phase noise is by means of phase locked loop techniques. The phase of the local oscillator must be corrected continuously in order to track the slow random phase drift of the oscillator under test with respect to the reference oscillator. Demodulation techniques for the spectral characterization of oscillators are described in detail in Chapter 4.

3.2.8 Standard definition of noise processes

3.2.8.1 Amplitude and phase noise processes

The basic noise processes that characterise the oscillator are the amplitude and phase noise processes $\varepsilon(t)$ and $\phi(t)$ defined above in the phasor model of the oscillator. As mentioned above the amplitude and the phase noise processes in a low noise oscillator are orthogonal in the complex plane and they can be detected independently with a very good rejection of each other. Therefore for the purpose of frequency stability analysis the amplitude noise can be neglected and it is the phase noise process that is relevant. Caution should be exercised here. AM to PM conversion can take place in some equipment. Even linear band-pass filtering produces AM to PM conversion if the transfer function is not perfectly symmetric with respect to the carrier frequency [Bernier et al., 1985]. For a more detailed discussion refer to Chapter 4.

3.2.8.2 Time error process

The time error process $x(t)$ can be defined as the normalised phase difference accumulated between an oscillator used as a clock and a reference oscillator considered as the reference clock. It is related to the phase noise process by:

$$x(t) = \frac{\phi(t)}{2\pi\nu_0} \tag{3.12}$$

where the nominal frequency ν_0 is expressed in [Hz] and the time error in [s].

3.2.8.3 Instantaneous frequency process

The normalised instantaneous frequency deviation $y(t)$ is defined as the derivative of the time error process $x(t)$:

$$y(t) = \frac{\mathrm{d}x(t)}{\mathrm{d}t} = \frac{1}{2\pi\nu_0}\frac{\mathrm{d}\phi(t)}{\mathrm{d}t} \tag{3.13}$$

$y(t)$ has no units. It is a frequency "deviation" in the sense that being associated with the complex envelope it describes not the instantaneous frequency of the oscillator signal $s(t)$ but the frequency deviation with respect to the nominal carrier frequency ν_0.

The instantaneous frequency $\nu_0(t)$ in [Hz] is related to $y(t)$ by the following:

$$y(t) = \frac{\nu(t) - \nu_0}{\nu_0} \tag{3.14}$$

For the purpose of frequency stability analysis it is the normalised frequency deviation $y(t)$ that is the most relevant quantity. The normalised frequency deviation is conserved after frequency multiplication or division and it also enables the direct comparison between the frequency instability levels of two oscillators of different nominal carrier frequencies.

3.2.9 Multiplicative and additive noise

3.2.9.1 Multiplicative noise

When the oscillator signal is amplified, transmitted, multiplied or is subjected to any other kind of processing, it can be degraded by either or both multiplicative and additive noise.

Multiplicative noise is actually a random modulation process that can be represented in a form similar to the complex envelope of the oscillator signal itself.

Assume that the signal of the oscillator is amplified and that the amplifier introduces random AM and PM modulations of the signal. The random modulation mechanism may be, for example, flicker noise in the polarisation currents of the transistors which modulates the gain and the phase-shift at the carrier frequency. The modulation function associated with the multiplicative noise process has the form:

$$M(t) = 1 + {}_m(t) + \mathrm{j}\phi_m(t) \tag{3.15}$$

The complex envelope of the oscillator signal at the output of the amplifier is the complex envelope at the input multiplied by the modulation function. Assuming that the phase noise processes of both the original signal and of the amplifier are both low-noise processes,

$$E\{(\varphi_i(t))^2\} << 1 \text{ and } E\{(\phi_m(t))^2\} << 1$$

the complex envelope of the low-noise oscillator at the input is:

$$\gamma_i(t) = 1 + {}_i(t) + \mathrm{j}\phi_i(t) \tag{3.16}$$

and, neglecting quadratic terms, the complex envelope of the oscillator signal at the output is:

$$\gamma_0(t) = M(t)\gamma_i(t) = 1 + \varepsilon_i(t) + \varepsilon_m(t) + \mathrm{j}(\phi_i(t) + \phi_m(t)) \tag{3.17}$$

The above result shows that in the low phase noise case, where the amplitude and phase noise processes are orthogonal, the multiplicative amplitude and phase noise processes simply add to the original amplitude and phase noise processes.

3.2.9.2 Additive noise

The case of additive noise can be treated as follows: assume a band-limited white noise process $n(t)$ of single-sided power spectral density N_0. According to Rice's representation theorem [Papoulis, 1983; Wozencraft et al., 1965] any band-pass random process can be decomposed into an in-phase and quadrature components with respect to an arbitrary carrier frequency ν_0:

$$n(t) = \sqrt{2}(n_p(t) + \mathrm{j}n_q(t))\exp(\mathrm{j}2\pi\nu_0 t) \tag{3.18}$$

Assuming that the band-pass process is a band limited white noise process of bandwidth 2B and single-sided spectral density N_0 and assuming, moreover, that its power spectral density is symmetrical with respect to the arbitrary carrier frequency ν_0, it can be shown that $n_p(t)$ y $n_q(t)$ are statistically independent low-pass white noise processes of single sided power spectral density N_0 and of bandwidth B [Wozencraft et al., 1965].

Rice's decomposition is in fact an analytic signal representation. Normalising with respect to the carrier peak amplitude A of the oscillator signal one obtains:

$$n(t) = A\left(\frac{\sqrt{2}}{A}n_p(t) + \mathrm{j}\frac{\sqrt{2}}{A}n_q(t)\right)\exp\left(\mathrm{j}2\pi v_0 t\right) \tag{3.19}$$

Comparing term to term with the analytic signal (3.4) of the low-noise oscillator, the contributions to amplitude and phase noise produced by additive noise are found to be two identical low-pass white noise processes of bandwidth B.

$$S_{\varepsilon\varepsilon}^{+}(f)\,|_{additive} = \mathrm{S}_{\phi\phi}^{+}(f)\,|_{additive} = \frac{N_0}{\frac{1}{2}A^2} = \frac{1}{S/N\,(1\,\mathrm{Hz})} \tag{3.20}$$

The one-sided power spectral density of both processes is equal to the inverse of the signal to noise ratio in a one Hertz bandwidth. Here the noise power in a one Hz bandwidth is defined as the power spectral density N_0 of the additive noise $n(t)$ and the signal power is defined as the power of the sinusoidal carrier of peak amplitude A.

3.2.10 Polynomial model

The conventional way of modelling the noise of oscillators is by the use of the polynomial model of $S_{yy}^{+}(f)$.

$$S_{yy}^{+}(f) = \sum_{\alpha=-2}^{\alpha=2} h_a f^{\alpha} \tag{3.21}$$

The phase/time process $x(t)$ being the integral of the frequency process $y(t)$ there is a direct correspondence between the power spectral densities of $x(t)$ and of $y(t)$:

$$S_{xx}^{+}(f) = \frac{S_{yy}^{+}}{(2\pi f)^2} = \frac{1}{(2\pi)^2}\sum_{\beta=-4}^{0} h_a f^{\beta} \quad \text{where} \quad \beta = \alpha - 2 \tag{3.22}$$

Table 3.1 summarises the different types of noise processes encountered in the polynomial model of oscillators. The model is convenient for classifying the different types of noise processes and for the translation of one type of frequency stability measurement into another.

Note that, in general, for a power law process $z(t)$ of power spectral density $S_{zz}^{+}(f) = k\, f^{\theta}$, the process is non-stationary when $\theta \le -1$ in the sense that the mean squared value of the process is given by:

$$E\left\{z^2(t)\right\} = R_{xx}(0) = \int_0^{\infty} S_{xx}^{+}(f)\quad df \tag{3.23}$$

which is undefined when the integral diverges at the origin. Note that here we use the generic process $z(t)$ and the exponent θ in order to state general properties that apply equally well to the time/phase process $x(t)$ and to the frequency process $y(t)$ with their respective exponents β and α.

The integral (3.23) does not converge at high Fourier frequencies for $\theta \geq -1$ because pure power-law processes are not band limited. This problem is only an artefact of the model, though, since the actual phase and frequency noise processes of oscillators are always band limited.

TABLE 3.1

Classification of power-law processes in the polynomial model

	Exponent α of $S_{yy}^{+}(f)$	Exponent β of $S_{xx}^{+}(f)$
White noise PM	2	0
Flicker noise PM	1	–1
White noise FM	0	–2
Flicker noise FM	–1	–3
Random walk FM	–2	–4

While f^{-1} is often called "flicker noise", a single integration of a white noise process yields an f^{-2} power law process, which is called a "random walk" or a one dimensional Brownian motion process.

[Greenhall, 1983] has shown that the power-law random processes found in oscillators can be characterised by power-spectral densities even when they are non-stationary ($\theta \leq -1$).

3.3 Characterisation: definitions and methods

3.3.1 Spectral domain

3.3.1.1 Basic definitions

In the spectral domain four measures are typically used to describe the oscillator: the power spectral density of $y(t)$, the power spectral density of $x(t)$, the power spectral density of $\phi(t)$ and $\mathcal{L}(f)$ a spectral measurement related to the concept of spectral purity.

The basic spectra are summarised in Table 3.2. The question whether the quantity of interest is $y(t)$, $x(t)$ or $\phi(t)$ depends upon the particular application. Most often, however, spectral measurements are used to characterise the spectral purity of the oscillator; therefore it is the spectral density of $\phi(t)$ that is usually given.

TABLE 3.2

Basic measurements in the spectral domain

Measurement	Symbol
PSD of $y(t)$	$S_{yy}^{+}(f)\ \left(\text{Hz}^{-1}\right)$
PSD of $x(t)$	$S_{xx}^{+}(f)\ \left(\text{s}^{2}\ \text{Hz}^{-1}\right)$
PSD of $\phi(t)$	$S_{ff}^{+}(f)\left(\text{rad}^{2}\ \text{Hz}^{-1}\right)$
Script $L(f)$	$\mathcal{L}(f)\quad\left(\text{Hz}^{-1}\right)$

The spectral purity measure $\mathcal{L}(f)$ is defined as:

$$\mathcal{L}(f) = \frac{1}{2} S_{\phi\phi}^{+}(f) \tag{3.24}$$

The relationship between the two measurements is justified below.

3.3.1.2 Spectral purity concepts

The spectral purity of an oscillator refers to the noise found on each side of the carrier when the signal of an oscillator is measured directly on a spectrum analyser.

Recalling that the power spectral density of the signal $s(t)$ of a low phase noise oscillator is given by (3.7), assume that the power spectral density of the oscillator signal is measured, normalised by the carrier average power $\frac{1}{2}A^2$ and translated from ν_0 to the origin into a low-pass spectrum.

$$\mathcal{L}(f) = \frac{S_{ss}^{+}(f+\nu_0)}{\frac{1}{2}A^2} = \delta(f) + S_{\varepsilon\varepsilon}(f) + S_{\phi\phi}(f) \tag{3.25}$$

This is to satisfy the definition of the spectral purity measure $\mathcal{L}(f)$ if the amplitude noise $\varepsilon(t)$ can be made negligible with respect to the phase noise $\phi(t)$. Then the spectral purity measure becomes identical to a measure of phase noise for all Fourier frequencies except the origin. This is often not true for frequency synthesisers and other devices where the oscillator signal has passed through a number of stages of amplification.

Note, however, that commercial spectrum analysers do not have the necessary resolution nor the dynamic range for the characterisation of the phase noise by a direct measurement of the band-pass signal $s(t)$. Moreover, the amplitude noise is often not negligible. In other words, the actual measurement of $\mathcal{L}(f)$ is not feasible in practice. Therefore, although commercial oscillators are almost always specified by $\mathcal{L}(f)$ the actual measurement involves the demodulation of $\phi(t)$ by phase locked loop techniques and the evaluation of its power spectral density using a digital low-pass spectrum analyser. The measurement of $S_{ff}^{+}(f)$ is then divided by a factor of two and labelled $\mathcal{L}(f)$.

3.3.2 Time domain

3.3.2.1 Introduction

Time-domain, in contrast to frequency-domain, deals with the effects of averaging over different values of time, τ. For useful time-domain measures, Fourier transform relationships exist between the Fourier frequency, f, and the time domain parameter, τ. Observing the dependence of a variance as τ is varied is often very insightful as to the oscillator characteristics.

Time-domain characterisation conventionally uses statistics computed from a discrete time series of average frequency samples obtained from a digital counter which is used to measure the beat note between the oscillator under test and a reference oscillator. The same discrete time series can, of course, be used in an FFT, for example, to compute the frequency domain-spectrum.

The approach of this presentation, on the other hand, consists of defining time domain variances in terms of the mean squared value of the continuous stationary random process obtained by the application of a linear operator, specific to the time domain variance of interest, to the instantaneous frequency process $y(t)$. Since $y(t)$ averaged over τ is $(x(t) - x(t-\tau))/\tau$, these variances can also be expressed in terms of $x(t)$.

This approach enables the establishment of the different variances in a simpler and more intuitive way than heretofore. The fact that, in practice, the variances are estimated over a finite number of samples taken from the underlying continuous process only affects the uncertainty of the estimation. The problem of the uncertainty on the practical estimation of the variances can be treated separately from the definition of the variance themselves and is already extesively investigated in the literature.

3.3.2.2 Basic concepts

3.3.2.2.1 Model of the frequency counter

Assume that an Oscillator Under Test (OUT) is to be characterised in the time domain. The conventional method is to heterodyne the original bandpass signal down to an arbitrary low beat frequency ν_0 by mixing the OUT signal with the signal of a reference oscillator much more stable than the OUT. The low frequency beat signal reproduces faithfully the original frequency fluctuations of the OUT, which is typically measured with a digital frequency counter.

The usual digital counter measures the frequency beat signal by counting the zero crossings of the signal. Zero crossing typically means the time-point where the signal voltage transverses positively through zero volts. For every sampling period, T, the counter delivers a sample ν_k which is the instantaneous frequency $\nu_b(t)$ [Hz] averaged over the gate opening time, τ.

In most frequency counters the sampling period, T, is larger than the averaging time, τ. There is therefore a dead-time, $T-\tau$ during which the counter computes the average frequency, resets the internal counters/registers, and waits for the next zero crossing of the signal. The next count can start only at the next zero crossing of the signal therefore in a frequency counter the dead-time is at least one period of the signal.

The dead time can be important as will be explained below because it biases the statistics when the samples are used for time domain measurements.

If the nominal average frequency ν_b is subtracted from the samples ν_k and if they are normalised with respect to the OUT mean frequency ν_0 one obtains the normalised samples y_k.

$$y_k = \frac{\nu_k - \nu_b}{\nu_0} \tag{3.26}$$

The discrete random process, or time series, $y_0, y_1, y_2, \ldots y_k$ coming from the frequency counter is therefore equal to the continuous random process $y(t, \tau)$, i.e. the moving average of $y(t)$ over τ as defined in next section, sampled with a period, T.

The time domain characterisation methods of the frequency stability are usually defined in terms of statistics performed over the discrete process $y_0, y_1, y_2, \ldots y_k$ coming from the frequency counter. We will show, however, that they are more easily understood when they are also defined in terms of the underlying continuous random processes.

3.3.2.2.2 Moving average operator

As mentioned above time domain variances are traditionally defined in terms of statistics performed on the discrete time series $y_0, y_1, y_2, \ldots y_k$ coming from the frequency counter. The samples obtained from the counter are a periodic sampling of the instantaneous frequency averaged over the gate opening interval, τ.

The frequency samples from the counter may be considered as discrete samples of the moving average of $y(t)$:

$$y_0, y_1, y_2, \ldots y_k = y(t_0, \tau), y(t_0 + T, \tau), y(t_0 + 2T, \tau), \ldots y(t_0 + kT, \tau)$$

where T is the sampling interval and where:

$$\mathrm{ma}(\tau)\{y(t)\} = y(t, \tau) = \frac{1}{\tau}\int_{t-\tau}^{t} y(\alpha)\, d\alpha \tag{3.27}$$

defines the causal moving average operator. Note that we mean by "causal" the fact that the operator applied at time t is defined only in terms of the present and past values of the process. The impulse response of the moving average operator ma(τ){} is a positive pulse function of duration and of amplitude 1/τ. The squared modulus of its transfer function is:

$$\left|H_{ma}(f)\right|^2 = \frac{\sin^2(\pi f\tau)}{(\pi f\tau)^2} \tag{3.28}$$

The moving average operator describes the averaging process taking place inside the frequency counter and will be used below for the analysis of the time domain variances.

3.3.2.2.3 Increment operator

The increment operator $\Delta(\tau)\{\}$ is defined as follows:

$$\Delta(\tau)\{x(t)\} = x(t) - x(t-\tau) \tag{3.29}$$

It can be applied recursively. When it is applied once the result is called the "first difference". When it is applied twice:

$$\Delta^{(2)}(\tau)\{(x(t)\} = \Delta(\tau)\{\Delta(\tau)\{x(t)\}\} = x(t) - 2x(t-\tau) + x(t-2\tau) \tag{3.30}$$

the result is called the "second difference", and so on.

The squared modulus of the transfer function associated with the increment operator $\Delta(\tau)\{\}$ is

$$\left|H_{\Delta}(f)\right|^2 = 4\sin^2(\pi f\tau) \tag{3.31}$$

Note that the transfer function of the increment operator is equivalent to a first-derivative operation at low Fourier frequencies close to the origin:

$$\left|H(f)\right|^2 = 4\sin^2(\pi f\tau) = 4(\pi f\tau)^2 \quad \text{for} \quad f << \frac{1}{\pi\tau} \tag{3.32}$$

As a consequence the first difference $\Delta(\tau)\{z(t)\}$ of a power-law process $z(t)$, of a power spectral density $S_{zz}^{+}(f) = k\,f^{\theta}$, yields a process of power spectral density $4k\sin^2(\pi f\tau)f^{\theta}$ that behaves like a power-law process of exponent $\theta + 2$ in the region of low Fourier frequencies close to the origin.

Therefore, for instance, the increment operator will transform a non-stationary flicker power-law process with exponent $\theta = -1$ into a stationary process with $\theta = +1$ behaviour in the region of low Fourier frequencies.

The resulting process is no longer a "power-law" process because the increment operator is not equivalent to the first-derivative operator for all Fourier frequencies but nevertheless it is stationary.

The capability of the increment operator, applied once or several times, to transform a non-stationary power-law process into a stationary process by its high-pass filtering action is fundamental in the definition of time domain frequency stability measurements.

3.3.2.3 Basic time domain measurements

3.3.2.3.1 The true variance

At first sight the logical way to characterise frequency stability of the OUT would be to evaluate the true variance of the frequency samples:

$$I^2(\tau) = E\{y_k^2\} \tag{3.33}$$

The time series y_k is a sampling of the continuous moving average process $y(t, \tau)$; therefore the mean squared value of the time series is equal to the mean squared value of the $y(t, \tau)$. Using (3.23) (3.28) and (3.A4) one obtains immediately:

$$I^2(\tau) = \int_0^\infty S_{yy}^+(f) \frac{\sin^2(\pi f \tau)}{(\pi f \tau)^2} \; df \tag{3.34}$$

which relates the true variance to the power spectral density of $y(t)$.

Unfortunately the integral (3.34) is infinite for non-stationary power-law noise processes, i.e., for $\alpha \le -1$.

In other words the time series y_k coming from the counter is non-stationary since the underlying continuous process $y(t,\tau)$ is non-stationary. The mean squared value of $y(t, \tau)$, which defines the true variance, diverges, i.e., becomes infinite when it is averaged over an infinite time because of the non-stationarity of the models representing the oscillators being characterised.

It is this observation, made in the early days of atomic frequency standards, that led to the definition of other, more appropriate, measures of frequency stability in the time domain.

3.3.2.3.2 The Allan variance

The traditional or classical Allan variance is defined as the average two-sample variance of the frequency samples y_k measured without a dead-time, i.e. with $T = \tau$.

$$\sigma_y^2(\tau) = \frac{1}{2} E\{(y_k - y_{k-1})^2\} \tag{3.35}$$

Observing that the action of taking the difference of two adjacent samples measured without a dead-time is equivalent to a first-difference operation applied over τ it follows that the Allan variance can be defined as the mean squared value of the continuous process $\Delta(\tau)\{y(t, \tau)\}$.

$$\sigma_y^2(\tau) = \frac{1}{2} E\{(\Delta(\tau)\{y(t,\tau)\})^2\} \tag{3.36}$$

In terms of linear operators the original process $y(t)$ is filtered by a moving average operator, modelling the action of the counter, followed by a first difference operator. Using (3.23), (3.28), (3.31) and (3.A4) one obtains immediately:

$$\sigma_y^2(\tau) = 2\int_0^\infty S_{yy}^+(f) \frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2} \; df \tag{3.37}$$

The above integral is defined and finite for all the power-law processes of Table 3.2. The process $\Delta(\tau)\{y(t, \tau)\}$, whose mean squared value defines the classical Allan variance, is indeed stationary although the original unfiltered process $y(t)$ was non-stationary. This comes from the high-pass filtering action of the first-difference operator.

If the counter producing the samples y_k has a dead time, then the moving average operation is taken over the gate opening interval τ and the first difference operation is taken over the sampling interval T. Therefore, the same definition (3.35) applied to the frequency samples y_k now yields a different result since by applying again (3.23), (3.28) (3.31) and (3.A4) as before we obtain:

$$\sigma_y^2(\tau,T) = 2\int_0^\infty S_{yy}^+(f) \frac{\sin^2(\pi f T)\sin^2(\pi f \tau)}{(\pi f \tau)^2} \; df \tag{3.38}$$

instead of (3.37).

The presence of a dead-time in the sampling of the average frequency introduces a bias in the statistics and does no longer yield the true Allan variance when the definition (3.35) is applied to the samples. The bias can be small in some cases. In the case of white noise FM, it is zero. In general it needs to be considered if it is intrinsic to a measurement system and cannot be avoided. See section 3.3.2.4.2.

3.3.2.3.3 The modified Allan variance

The modified Allan variance was originally defined in terms of the samples x_k of the phase/time process $x(t)$ [Allan et al., 1981]:

$$\mathrm{mod}\,\sigma_y^2(n\tau_0) = \frac{1}{2\,(n\tau_0)}\,\frac{1}{2}\,E\left\{\left(\frac{1}{n}\sum(x_{i+2n} - 2x_{i+n} + x_i)\right)^2\right\} \tag{3.39}$$

In order to make easier a comparison with the classical Allan variance, the definition (3.39) can be rewritten in terms of frequency samples:

$$\mathrm{mod}\,\sigma_y^2(n\tau_0) = \frac{1}{2}\,E\left\{\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{n}\sum_{k=1}^{n} y_{i+k+n} - \frac{1}{n}\sum_{k=1}^{n} y_{i+k}\right)\right]^2\right\} \tag{3.40}$$

where the y_k are frequency samples measured without a dead time and averaged over the sampling interval τ_0.

This new formulation (3.40) can be interpreted as follows. The inner summations group the elementary samples n by n into super samples averaged over $n\tau_0 = \tau$. The outer summation, on the other hand, is a discrete moving average (dma) operation, performed over the averaging interval $n\tau_0 = \tau$, and applied to the first difference of the super samples.

Therefore the modified Allan variance can be defined very simply in terms of the underlying continuous processes as follows:

$$\mathrm{mod}\,\sigma_y^2(\tau) = \frac{1}{2}\,E\left\{\left[dma(n,\tau_0)\{\Delta(\tau)\{y(t,\tau)\}\}\right]^2\right\} \tag{3.41}$$

where the discrete moving average operator applied over the averaging interval $n\tau_0 = \tau$ is defined as:

$$dma\,(n,\tau_0)\{z(t)\} = \frac{1}{n}\sum_{i=0}^{n-1} z(t - i\,\tau_0) \tag{3.42}$$

The squared modulus of the transfer function associated with the discrete moving average operator is:

$$\left|H_{dma}(f)\right|^2 = \frac{\sin^2(\pi f \tau)}{n^2 \sin^2\left(\frac{(\pi f \tau)}{n}\right)} \tag{3.43}$$

where the averaging interval τ is n times the elementary sampling interval τ_0.

This last result shows that the modified Allan variance is identical to the classical Allan variance except for a supplementary discrete moving average operation.

Using the transfer functions (3.23), (3.28), (3.31) and (3.A4) we obtain immediately:

$$\mathrm{mod}\,\sigma_y^2(n,\tau) = 2\int S_{yy}^{+}(f)\,\frac{\sin^2(\pi f \tau)}{(n\pi f\tau)^2 \sin^2\left(\frac{(\pi f \tau)}{n}\right)}\,df \tag{3.44}$$

This is the expression generally found in the literature. The advantage of the modified over the classical Allan variance is the fact that it produces distinct slopes as a function of τ for each of the power law noise processes defined in the polynomial model. The disadvantage, on the other hand, is the fact that the modified Allan variance depends not only on the averaging time τ but also on the number n of elementary samples used to build the interval τ and, indirectly, also on the sampling interval τ_0.

This is simply due to the fact that, like in any digital processing operation, aliasing occurs when the sampling rate does not satisfy the condition of the sampling theorem, in other words when the sampling rate is smaller than the Nyquist frequency which is defined as twice the bandwidth of the signal.

Assuming, on the other hand, that the sampling rate $1/\tau_0$ is higher than twice the bandwidth B of the process y(t), then the sampling theorem is satisfied and the discrete moving average operator yields approximately the same result as the continuous moving average operator.

It can be verified that when $n \to \infty$ for constant τ, which is equivalent to allowing the sampling rate $1/\tau_0$ to tend to ∞, the discrete transfer function (3.43) tends to the continuous transfer function (3.34).

Therefore, by applying the condition of the sampling theorem, i.e. for $1/\tau_0 > 2B$, where B is the bandwidth of $y(t)$, an analogue of the modified Allan variance can be established using the continuous moving average operator (3.27) instead of the discrete moving average operator (3.42), yielding:

$$\mathrm{cmod}\,\sigma_y^2(\tau) = \frac{1}{2} E\left\{\left[\Delta(\tau)\left\{ma^{(2)}(\tau)\,\{y(t)\}\right\}\right]^2\right\} \tag{3.45}$$

instead of (3.41). Similarly the modified Allan variance can be established in the spectral domain using the transfer function of the continuous moving average operator (3.28) instead of the transfer function of the discrete moving average operator (3.43), yielding:

$$\mathrm{cmod}\,\sigma_y^2(\tau) = 2\int_0^\infty S_{yy}^+(f)\frac{\sin^6(\pi f \tau)}{(\pi f \tau)^4}\; df \tag{3.46}$$

instead of (3.44).

The “continuous” modified Allan variance, noted cmod, is the limit of the modified Allan variance measured without truncation, (see section on truncation below), and in compliance with the sampling theorem.

The results about the cmod Allan variance summarised in Tables 3.5 to 3.7 were computed analytically from (3.46) and establish that the cmod variance has distinct slopes for different power-law processes and, at the same time, is independent of n, of τ_0 and of the system bandwidth B as shown and verified in [Bernier, 1987].

The cmod variance therefore corrects the two drawbacks of the classical Allan variance: the fact that the slope of the classical variance is the same for both white PM and flicker PM and the dependence on the system bandwidth.

Moreover the results summarised in Tables 3.5 to 3.7 also show that even if the condition of the sampling theorem is not satisfied, the conventional modified Allan variances has the same properties as the cmod Allan variance for all noise processes except white PM on the condition that $n \gg 1$. In the case of white PM the dependence of the modified Allan variance on the system bandwidth is the same as that of the classical Allan variance.

After this important conclusion, let us define yet another formulation of the modified Allan variance, which does not bring new insight, but that is very convenient for the efficient practical estimation of the modified Allan variance.

Define $w(t)$ as the integral of $x(t)$:

$$w(t) = \int_{t_0}^{t} x(\alpha)\,\mathrm{d}\alpha \tag{3.47}$$

Now, by the use of (3.47), the moving average of $x(t)$ can be expressed as:

$$x(t,\tau)=\frac{1}{\tau}\Delta(\tau)\{w(t)\} \tag{3.48}$$

By applying (3.47) a second time and by taking the first difference one obtains the continuous modified (cmod) Allan variance, i.e. the modified Allan variance in which the supplementary moving average operator is a continuous operator, yielding:

$$\text{cmod}\,\sigma_y^2(\tau)=\frac{1}{2}E\left\{\left(\frac{1}{\tau^2}\Delta^{(3)}(\tau)\{w(t)\}\right)^2\right\} \tag{3.49}$$

This is the continuous case. In the discrete case, on the other hand, when the samples come out from the counter with a sampling interval τ_0, if one defines w_k as the discrete sum of all the past x samples,

$$w_k=\sum_{i=0}^{k}x_i \tag{3.50}$$

then the discrete moving average required for the computation of the discrete modified Allan variance is given by:

$$\frac{1}{n}\sum_{i=0}^{n-1}x_{k-i}=\frac{1}{n}(w_k-w_{k-n}) \tag{3.51}$$

As a consequence the discrete modified Allan variance can be defined in terms of the third difference of w(t):

$$\text{mod}\,\sigma_y^2(n\tau_0)=\frac{1}{2}\frac{1}{(n^4\tau_0^2)}E\{(w_k-3w_{k-n}+3w_{k-2n}-w_{k-3n})^2\} \tag{3.52}$$

This last result can be used for the efficient estimation of the modified Allan variance as shown in the section about algorithms below. See Chapter 4 and Table 4.7 for confidence intervals for $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$.

3.3.2.3.4 The time interval error

The time interval error (TIE) is a notion used in the predication of time scales and also in telecommunications for the synchronisation of digital networks [Kartaschoff, 1987]. It constitutes a good example of a direct application of the Allan variance.

Assume that at time t two oscillators are both synchronised (time difference set to zero) and syntonised (frequency offset set to zero). The time interval error $TIE(t,\tau)$ is the time error $x(t)$ accumulated between the two oscillators, considered as clocks, at a time $t+\tau$ in the future.

There are different possible ways of defining the TIE depending upon which estimator of the initial frequency offset is chosen. And, as a consequence, the statistical properties of the process $TIE(t,\tau)$ also depend upon which estimator is chosen.

Suppose the frequency offset is estimated by averaging the instantaneous frequency $y(t)$ during the interval τ preceding the instant t. The frequency estimator at time t is therefore $y(t,\tau)$ and the TIE becomes:

$$TIE(t,\tau)=x(t+\tau)-x(t)-\tau y(t,\tau) \tag{3.53}$$

The first term of (3.53) is the free running non-stationary time error random process $x(t)$ at the time $t+\tau$ in the future. The second term is the initial time difference $x(t)$. Subtracting it is equivalent to synchronising the oscillators at time t. The third term is the time difference accumulated after an interval τ because of the initial frequency offset $y(t,\tau)$. Subtracting it is equivalent to syntonising the oscillators at time t.

It can be shown that the TIE defined by (3.53) is identical to the second difference of $x(t)$:

$$TIE(t,\tau) = x(t+\tau) - x(t) - \Delta(\tau)\{x(t)\} = \Delta^{(2)}(\tau)\{x(t+\tau)\} \tag{3.54}$$

which is stationary for all the power-law processes found in the polynomial model. Therefore although the exact behaviour of the TIE cannot be predicted, since it is a random process, its mean squared value can be computed.

The mean squared value of the TIE is equal to the mean square value of the second difference of $x(t)$ which can be expressed as a simple function of the classical Allan variance:

$$E\{TIE^2(t,\tau)\} = 2\tau^2\sigma_y^2(\tau) \tag{3.55}$$

The mean squared value of the second difference of $x(t)$ is in fact the second structure function of $x(t)$. Structure functions are discussed in the section below about other time domain variances.

Equation 3.55 does not imply that optimum prediction procedures were used – only a particular procedure. If a clock has white FM noise and optimum prediction is used, then the TIE is given by $\tau\sigma_y(\tau)$. In this case the optimum frequency estimate is the mean frequency from the infinite past.

A generalisation of the TIE is reported in [Bernier, 1988] in which a predictor of order n of $x(t)$ is defined for which the mean squared value of the error on the prediction is given by the structure function of order $n + 1$. For example if the TIE corrects only the effect of the initial frequency offset as in (3.54) the predictor is of order 1 and the error on the prediction is the structure function of order 2 given by (3.55). If the predictor corrects not only for the initial frequency offset but also for the initial frequency drift, then the predictor is of order 2 and the error on the prediction is the structure function of order 3 and so on. Prediction modelisation of time-scales is an important topic and is treated extensively in the literature [Tavella et al., 1991; Allan, 1987].

3.3.2.3.5 The time variance

The time variance $\sigma^2_x(\tau)$ [Allan et al., 1991] is defined as:

$$\sigma_x^2(\tau) = \frac{1}{3}\tau^2 \bmod \sigma_y^2(\tau) = \frac{1}{6}E\{\Delta^{(2)}(\tau)\{\bar{x}(t)\}\}^2 \tag{3.56}$$

where the $\bar{x}(t)$ is the average time error over an interval τ ending at t.

The time variance $\sigma_y^2(\tau)$ is now used as a standard in the telecommunications industry for specifying network performance, etc. It is a kind of TIE mean squared value estimator similar to the one defined in previous section. More detailed properties of the time variance can be found in [NIST, 1990].

3.3.2.3.6 Other time domain measurements

Other time domain measurements are described in the literature: structure functions [Lindsey et al., 1976; Greenhall, 1983], the Hadamard variance [Rutman, 1978], the High-Pass variance [Rutman, 1978], etc.

As shown above for the Allan variance, the principle of time domain variances is always the same: the variance is defined as the mean squared value of a stationary process produced by the action of a linear operator applied to the original non-stationary low-pass frequency process $y(t)$ or phase process $x(t)$.

In particular the structure function of order n of the process $z(t)$ is defined as the mean squared value of the increment of order n of a the process [Lindsey et al., 1976]:

$$D_z^{(n)}(t,\tau) = E\{(\Delta^{(n)}(\tau)\{z(t)\})^2\} \tag{3.57}$$

For example assume a random walk FM noise process. The power spectral density of the frequency $y(t)$ is a power-law of the Fourier frequency f with exponent –2 and the power spectral density of the time function $x(t)$ is a power law with exponent –4. As shown by equation (3.32) each application of the increment operator increases the exponent by +2 at Fourier frequencies close to zero. Therefore the first increment of $y(t)$ is stationary and the second increment of $x(t)$ is also stationary. The structure functions $D_y^{(1)}(t,\tau)$ and

$D_x^{(2)}(t,\tau)$ are therefore independent of t and constitute time domain variances i.e. mean squared values depending on the time difference parameter τ like the classical Allan variance. In fact the Allan variance is directly related to the second structure function of $x(t)$:

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2} D_x^{(2)}(\tau) \tag{3.58}$$

which explains the result (3.55).

3.3.2.3.7 The multi-variance analysis

The multi-variance analysis should be considered as an estimator of the PSD of $y(t)$ rather than a classical variance measurement. Since the PSD of $y(t)$ may be modelled as a sum of power laws (from f^{-2} to f^{+2} frequency noise), the goal of this method is to measure each noise level. The principle consists in using several variances over the same signal. These different variances must be chosen in order to get an accurate measurement for low frequency noises (Allan variance, Picinbono variance) as well as for high frequency noises (modified Allan variance, Time variance). Moreover, the drifts may also be measured by this method.

The relevant parameters (noise levels and drift coefficients) are determined by a weighted least square method: knowing the response of each variance for each type of noise (see Tables 3.8 and 3.9), and weighting each variance measurement by the inverse of its confidence interval, the noise levels and drift coefficients are calculated as the most probable parameters in the sense of the least squares.

This method is more sensitive than a single variance measurement since the accuracy of the measurement is at least as good as the most accurate variance of the set of variances. The choice of the variance thus determines the sensitivity of the multi-variance analysis: the best set of variance would be a set containing variances specially designed for each relevant parameter. But this choice may be adapted to each case [Vernotte et al., 1995].

The second improvement of this method concerns the dynamics. The dynamics may be defined as the range in which a noise level or a drift coefficient can vary without swamping the other ones or being swamped by the other ones. The dynamics depends on the number and on the type of the variances chosen, but is always greater than the dynamics of each variance of the set.

The main advantage of the multi-variance analysis is the determination of the uncertainty domain over each parameter measurement. The confidence intervals are estimated from the dispersion of the variance results by using a singular value decomposition. Moreover, since the assessment of the precision leads to associate an eigen vector to each noise level or drift coefficient, it is possible to calculate the angle between each eigen vector. A small angle between two eigen vectors implies that the parameters corresponding to these vectors yield similar effects over the variance responses. In this connection, the separability may be defined as the aptitude of a variance or a set of variances to distinguish between two noises. Thus, the determination of the angles between the eigen vectors corresponding to the parameters yields a quantitative estimation of the separability of the multi-variance analysis and provides a criterion to estimate the measurement reliability. Only the types of noises and drifts that do exist in the signal are then measured. For a full description and measurement results of the multi-variance analysis see [Vernotte et al., 1993] and [Walter, 1992].

3.3.2.4 Pitfalls

3.3.2.4.1 Effect of zero-crossing detection

Time domain measurements are estimated using either phase or frequency samples coming from a digital counter. The counter detects the zero-crossings of the signal which is usually a beat note between the OUT and a reference oscillator.

For the measurement of precision oscillators it is usually necessary to use a dedicated zero-crossing detector between the signal and the input of the counter.

The reason is that if the slew rate of the signal at the input of the counter is not high enough, the noise in the trigger circuitry of the counter will contribute to the measured Allan variance. Note that, paradoxically, the higher the time resolution of a counter is, the higher the bandwidth of the trigger circuitry and the higher the RMS amplitude of the additive noise is in the trigger stage of the counter.

The design of the dedicated zero-crossing detector must be adapted to the characteristics of the signal to be measured (amplitude bandwidth, beat note frequency etc.). If the transitions at the output of the zero-crossing detector are sharp enough, then the noise of the trigger circuitry in the counter becomes negligible.

However it is then the noise contribution of the zero-crossing detector itself that determines the noise floor of the measuring system. The noise due to zero-crossing detection is an additive wideband noise. Therefore it appears as band-limited white phase noise in the Allan variance measurement and it is characterised by a τ^{-1} behaviour of the Allan deviation $\sigma_y(t)$.

3.3.2.4.2 Effect of dead-time

As mentioned above the Allan variance can be defined by (3.35) in terms of frequency samples only if there is no dead-time in the sequence of measurements.

The bias in the Allan variance measurements introduced by the dead-time was studied extensively and Tables can be found in the literature [Barnes et al., 1989; Lesage et al., 1979; NIST, 1990]. As mentioned above in the section on the Allan variance, however, the availability of modern time-tagging counters that can sample $x(t)$ and the recent possibility of transforming a conventional time interval counter into a time-tagging counter by post-processing software and an auxiliary “picket fence” reference signal [Greenhall, 1989] make it easy to avoid a dead-time in the measurements.

Besides, grouping neighbouring frequency samples averaged over an interval τ_0 is equivalent to acquire one sample averaged over an interval $n\tau_0$ only if there is no dead-time in the sequence of measurements. The bias introduced by the grouping of samples in the presence of dead-time is analysed in [Lesage, 1983]. However the same remark as before applies: recent techniques make it easy to avoid a dead-time.

3.3.2.4.3 Effect of system bandwidth

The Allan variance transfer function in integral (3.37) goes to zero as f^{-2} for $f \to \infty$. Therefore for a pure power-law frequency process $y(t)$ with exponent $\alpha \geq 1$ the integral does not converge at high Fourier frequencies. Note that this happens only if the process $y(t)$ is a pure power-law process, i.e. if its bandwidth is infinite.

In a real oscillator the bandwidth of the process $y(t)$ is finite because the signal is band-limited and the actual upper bound of integral (3.37) is the bandwidth B of $y(t)$. Therefore the integral is always bounded on the high frequency side.

For $\alpha < 1$ the f^{-2} high-frequency cut-off behaviour of the transfer function is sufficient by itself to limit the bandwidth and to make the integral converge. The Allan variance is therefore independent of the bandwidth of the frequency process $y(t)$.

For $\alpha \geq 1$, on the other hand, the f^{-2} high-frequency cut-off behaviour of the kernel is not sufficient by itself to limit the bandwidth and to make the integral converge. Therefore the Allan variance is a function of the integral upper boundary, i.e. of the bandwidth B of $y(t)$.

The real-life power-law processes found in oscillators are always low-pass filtered and are well modelled by power-law behaviour only at low Fourier frequencies. For $\alpha \geq 1$, therefore, the Allan variance is a function not only of the bandwidth B but also of the specific shape of the low-pass filter used to limit the bandwidth. This has been taken into account in tables published in the literature [NIST, 1990].

Note, moreover, that the bandwidth B, as it appears in the Allan variance measurement, is usually limited by the low-pass bandwidth of the measuring system used after demodulation rather than by the bandwidth of the original band-pass oscillator signal.

The fact that the Allan variance of a white phase noise process is not only a function of the power spectral density of the white phase noise but also a function of the measuring system bandwidth constitutes a weakness of the classical Allan variance.

On the other hand the transfer function of the continuous modified Allan variance, as defined by (3.46), has a f^{-4} high-frequency cut-off behaviour because there are two cascaded moving average operators in its definition. In this case, then, the transfer function is sufficient to limit the bandwidth for all power-law processes in the polynomial model. The modified Allan variance is therefore independent of the bandwidth for all power law processes in the polynomial model if the condition of the sampling theorem is satisfied.

If the condition of the sampling theorem is not satisfied however, aliasing effects make the modified Allan variance dependent on the bandwidth in the case of white phase noise. For the other power-law processes, including flicker of phase, the high-frequency cut-off behaviour of the transfer function defined by (3.44) is sufficient for convergence.

3.3.2.4.4 Truncation effects

The main lobe of both the classical and of the modified Allan variance transfer functions is centred about the Fourier frequency:

$$\frac{1}{2\tau}$$

Therefore, if one attempts to compute the classical or modified Allan variance for averaging times too small i.e. for:

$$\tau < \frac{1}{2B}$$

it happens that the band-pass transfer function of the variance is centred on a Fourier frequency higher than the low-pass bandwidth of the signal $y(t)$.

Therefore the resulting Allan variance goes rapidly to zero as τ is further decreased. In this region of truncation the variance depends strongly on the shape of the low-pass filter used to limit the bandwidth of the signal and the value of the variance will be biased low.

Note also that, in the case of the modified Allan variance, if the signal is sampled in compliance with the sampling theorem we have:

$$\tau_0 < \frac{1}{2B}$$

Therefore truncation occurs for small values of n and disappears only for $n >> 1$ since $\tau = n\tau_0$ [Bernier, 1987].

3.3.2.5 Algorithms

3.3.2.5.1 Frequency averaging by phase sampling

Both the classical and the modified Allan variances can be defined either in terms of frequency samples or in terms of phase samples.

Because $y(t)$ is simply the derivative of $x(t)$ it is very easy to transform any definition that uses one quantity into an equivalent definition in terms of the other quantity.

The two following identities are very useful for the calculation of the time domain variances:

$$ma(\tau)\{y(t)\} = \frac{1}{\tau}\Delta(\tau)\{x(t)\} \tag{3.59}$$

$$\Delta(\tau)\{y(t)\} = \frac{1}{\tau}\Delta^{(2)}(\tau)\{x(t)\} \tag{3.60}$$

Assuming that the phase/time function $x(t)$ is sampled with a sample period τ, the time series

$$x_1, x_2, x_3, \ldots x_n$$

enables one to compute the series of frequency samples averaged over τ using the first difference of the x samples:

$$y_1, y_2, y_3, \ldots$$

where:

$$y_n = \frac{1}{\tau}(x_n - x_{n-1}) \tag{3.61}$$

The second difference of $x(t)$ yields the first difference of the frequency necessary for the computation of the Allan variance.

$$\Delta y_1,\ \Delta y_2,\ \Delta y_3 \ldots \Delta_{yn}$$

where:

$$\Delta y_n = y_n - y_{n-1} = \frac{1}{\tau}(x_n - 2x_{n-1} + x_{n-2}) \tag{3.62}$$

Regarding the hardware, averaging directly the frequency in a digital counter inevitably produces a dead-time since, after a sample is delivered, the counter must wait at least one period of the signal waiting for the next zero-crossing before the next count is started.

On the other hand if it is the time error function $x(t)$ that is sampled the above results show that the frequency samples averaged over τ without a dead-time can be computed easily, and the value of τ can be selected in the software only constrained by $\tau = n\tau_0$, where n is an integer.

In practice the time error function $x(t)$ can be sampled by time-tagging the zero-crossings of the signal using a “time-tagging” counter. If a time-tagging counter is not available, though, [Greenhall, 1989] has published an algorithm, called the “picket fence” algorithm, for computing directly samples of $x(t)$ using a conventional time interval counter.

3.3.2.5.2 Computation of classical Allan variance

The classical Allan variance can be estimated from N successive samples of $x(t)$, with sampling interval τ_0, using the following formula:

$$\sigma_y^2(\tau) = \frac{1}{2\tau^2(N-2n+1)}\sum_{i=0}^{N-2n}(x_{k-i} - 2x_{k-i-n} + x_{k-i-2n})^2, \qquad \text{where} \quad \tau = n\tau_0 \tag{3.63}$$

This is an “overlapping” estimate over N samples of (3.35) computed using the transformation (3.61). The formula is causal and therefore usable in real-time calculations.Assume that x_k is the last sample acquired, defining the present, then formula (3.63) involves only past samples in the calculation.

The uncertainty on the true mean square value due to the calculation of the estimation from a finite number of samples was determined in [Lesage et al., 1973] and [Lesage et al., 1976]. Lesage and Audoin results are summarised in [Allan et al., 1988] and [IEEE, 1988].

Note, however, that the reason why the variance is computed from discrete samples is an artefact coming from the use of a digital counter to collect the data. In general the Allan variance can be defined as the mean square value of the continuous process $\Delta(\tau)\{y(t, \tau)\}$ and it could be also estimated by taking the mean square value of the continuous process observed during a finite observation interval. The uncertainty on the estimation of the Allan variance from the mean squared value of $\Delta(\tau)\{y(t,\tau)\}$ taken over a finite time interval is treated in [Greenhall, 1983]. The effect of sampling the continuous process $\Delta(\tau)\{y(t,\tau)\}$ for the estimation of the Allan variance is treated in [Walter, 1994]. It yields a better confidence for the estimate.

The estimator (3.63) also has the property that for each value of τ only 5 registers are required in the computer in order to update the Allan variance in a continuous way (see Table 3.3).

TABLE 3.3

Registers allocation for the classical Allan variance computation

Register 1	x_i
Register 2	x_{i+n}
Register 3	x_{i+2n}
Register 4	Sum of variances
Register 5	Number of variances

Assume that the sampling period of $x(t)$ is τ_0. For each value of $\tau = n\tau_0$ for which one wants to compute the Allan variance, there are 5 registers to reserve.

From the stream of $x(t)$ samples coming every τ_0 from the counter one picks up one every n samples. The spacing between samples is therefore $\tau = n\tau_0$. From the three values spaced $n\tau_0$ apart, the squared second-difference is computed. Registers 4 and 5 are updated. The contents of register 1 are replaced with the next x_i spaced τ_0 later, etc. for 2 and 3, and registers 4 and 5 are updated again…

Note, finally, that in practice it is always the standard deviation, i.e. the square root of the variance, that is specified and not the variance itself. The practical consequence is that on a log-log plot the slopes associated with the standard deviation are half the slopes associated with the variance.

3.3.2.5.3 Computation of modified Allan variance

At first sight the computational load associated with the modified Allan variance is much larger, especially for large values of n, because of the summation over n in the definition (3.39). In fact, however, the computational load is essentially the same as for the classical variance except for one more register to reserve.

From the modified Allan variance formulation (3.52) an N samples overlapping estimate of the modified Allan variance can be defined:

$$\bmod \sigma_y^2(n\tau_0) = \frac{1}{2}\frac{1}{(n^4\tau_0^2)}\frac{1}{N-3n+1}\sum_{i=1}^{N-3n+1}(w_{k-i} - 3w_{k-i-n} + 3w_{k-i-2n} - w_{k-i-3n})^2 \tag{3.64}$$

where the averaging time τ is n times the elementary sampling interval τ_0.

This last result demonstrates that for each value of τ only 6 registers are required in the computer in order to update the modified Allan variance in a continuous way (see Table 3.4).

One common register is required to store w_k i.e. the sum of all the previous x_k coming out from the counter as defined by (3.50).

TABLE 3.4

Registers allocation for the modified Allan variance computation

Register 1	w_i
Register 2	w_{i+n}
Register 3	w_{i+2n}
Register 4	w_{i+3n}
Register 5	Sum of variances
Register 6	Number of variances

From the stream of w_k samples updated every τ_0 in the common register, one picks up one every n samples. As before the spacing between samples is $\tau = n\tau_0$. The same procedure is followed as in computing (3.63) only the third-difference squared is used instead.[Greenhall, 1992]

In conclusion we see that by defining the time series w_i, which is simply the sum of all the previous x_k, the computation effort to get the modified Allan variance is essentially the same as for the classical Allan variance. Software is available from NIST for computing these variances.

3.3.2.5.4 Summary

For computational purposes, Tables 3.5 and 3.6 give symmetrical formulae for the estimation of the classical and modified Allan variances on the basis of either frequency or phase elementary samples. The spectral formulae are also given on the basis of either the phase noise or frequency noise power spectral densities.

TABLE 3.5

Symmetrical formulae for classical Allan variance

$\sigma_y^2(t) = 2\int_0^\infty S_{yy}^+(f)\dfrac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}\,\mathrm{d}f$
$\sigma_y^2(t) = \dfrac{2}{\nu_0^2}\int_0^\infty S_{\varphi\varphi}^+(f)\dfrac{f^2\sin^4(\pi f\tau)}{(\pi f\tau)^2}\,\mathrm{d}f$
$\sigma_y^2(n\tau_0) = \dfrac{1}{2}\,\dfrac{1}{(n^4\tau_0^2)}\,\dfrac{1}{(N-2n+1)}\sum_{i=0}^{N-2n}(x_{k-i} - 2x_{k-i-n} + x_{k-i-2n})^2$
$\sigma_y^2(n\tau_0) = \dfrac{1}{2}\quad\dfrac{1}{(N-2n+1)}\sum_{i=0}^{N-2n}\left[\dfrac{1}{n}\sum_{j=0}^{n-1}y_{k-i-j} - \dfrac{1}{n}\sum_{j=0}^{n-1}y_{k-i-j-n}\right]^2$

TABLE 3.6

Symmetrical formulae for modified Allan variance

$\mathrm{mod}\ \sigma_y^2(n,\tau) = 2\int_0^\infty S_{yy}^+(f)\dfrac{\sin^6(\pi f\tau)}{(n\pi f\tau)^2\ \sin^2\left(\dfrac{\pi f\tau}{n}\right)}$
$\mathrm{mod}\ \sigma_y^2(n,\tau) = \dfrac{2}{\nu_0^2}\int_0^\infty S_{\varphi\varphi}^+(f)\dfrac{f^2\sin^6(\pi f\tau)}{(n\pi f\tau)^2\ \sin^2\left(\dfrac{\pi f\tau}{n}\right)}$
$\mathrm{mod}\ \sigma_y^2(n\tau_0) = \dfrac{1}{2}\ \dfrac{1}{\left(n^4\tau_0^2\right)}\ \dfrac{1}{(N-3n+1)}\sum_{i=0}^{N-3n}\left(w_{k-i} - 3w_{k-i-n} + 3w_{k-i-2n} - w_{k-i-3n}\right)^2$
$\mathrm{mod}\ \sigma_y^2(n\tau_0) = \dfrac{1}{2}\ \dfrac{1}{(N-3n+1)}\sum_{i=0}^{N-3n}\left[\dfrac{1}{n}\sum_{l=0}^{n-1}\left(\dfrac{1}{n}\sum_{j=0}^{n-1}y_{k-i-j-l} - \dfrac{1}{n}\sum_{j=0}^{n-1}y_{k-i-j-l-n}\right)\right]^2$

The relationship $\mathrm{S}_{yy}^+ = \dfrac{f^2}{\nu_0^2}\mathrm{S}_{\phi\phi}^+$ used in the transformation is explained simply by the fact that $y(t)$ is the derivative of $x(t)$.

The formulae found in Tables 3.5 and 3.6 are derived from (3.37), (3.44), (3.57) and (3.58). They are subject to the same restrictions and definitions. Although the variances can be estimated from elementary frequency samples averaged without a dead time over τ_0, it is important to note that the estimation of the variances using phase samples is less computationally expensive as explained in the text.

3.3.2.6 Applications

Originally the classical Allan variance was introduced as a means to characterise precision frequency sources at a time when spectral analysis at very low Fourier frequencies was not possible. Nowadays it is technically feasible to sample $x(t)$ at a sampling rate higher than the Nyquist rate and to compute numerically from the samples the power spectral density of $x(t)$ or of $y(t)$ at all Fourier frequencies. Given the power spectral density it is possible to identify the noise processes and to compute every kind of time domain variances.

Nevertheless time-domain variances are still useful tools to characterise oscillators and to identify the noise processes. One reason is the required acquisition and processing time for a given degree of confidence. On the one hand time-domain variances can be updated continuously as new data is acquired. On the other hand spectral analysis requires the sampling of many slices of data, with each slice having a duration equal to the inverse of the required resolution, before an accurate power spectral density estimation is obtained.

Another reason is that most practical measuring systems sample $y(t)$ or $x(t)$ without respecting the sampling theorem. In this case spectral analysis of the data would be ruined by aliasing while the classical Allan variance is still meaningful.

For the purpose of the identification of the noise processes defined in the polynomial model, the modified Allan variance is superior to the classical Allan variance because it yields different slopes for all the useful noise process models. This is to be compared to the classical Allan variance that yields the same τ^{-2} dependence for both white PM and flicker PM. Moreover, on the condition that the sampling theorem is satisfied, the modified Allan variance is independent of the measurement bandwidth B for all power-law processes while the classical Allan variance is a function of the bandwidth B for both white PM and flicker PM. In the particular case of flicker PM the modified Allan variance is independent of the system bandwidth even if the sampling theorem is not satisfied.

3.3.2.7 Conversion between time & frequency domains

Table 3.7 shows the correspondence between the log-log slopes associated with the spectral domain and time domain measurements of power-law processes.

TABLE 3.7

Correspondence of log-log slopes between spectral and time domains for power-law noise processes

	Slope of log-log plot			
	Spectral domain		Time domain	
	$S^{+}_{yy}(f)$	$S^{+}_{\phi\phi}(f)$	$\sigma^2_y(\tau)$	$\text{mod}\,\sigma^2_y(\tau)$
Slope	a	$b = a - 2$	m	m'
Random walk FM	–2	–4	1	1
Flicker FM	–1	–3	0	0
White FM	0	–2	–1	–1
Flicker PM	1	–1	–2	–2
White PM	2	0	–2	–3

As mentioned above the modified Allan variance yields different slopes for the different noise processes. This property facilitates the identification of the processes. This is to be compared with the classical Allan variance that yields the same τ^{-2} slope for both white PM and flicker PM.

On the other hand the classical Allan variance is a function of the bandwidth B for white PM and flicker PM. The reasons for this dependence are explained above. As a consequence the Allan variance for a white PM process allows the determination of the mean squared value N_0 B of the process and does not allow the determination of the spectral density N_0 unless the bandwidth B of the system is properly calibrated. On the other hand, Allan variance is independent of the measurement bandwidth B for all power-law processes on the condition that the sampling theorem is satisfied. The modified Allan variance measured under this condition is noted “cmod” in Table 3.8 and the coefficients A_c to E_c are to be used for the conversion (see Table 3.9).

TABLE 3.8

Conversion factors for translation of power-law processes

	$S^{+}_{yy}(f)$	$\sigma^2_y(\tau)$	mod $\sigma^2_y(\tau)$	cmod $\sigma^2_y(\tau)$
Random walk FM	$h_{-2} f^{-2}$	$A\, h_{-2}\, t$	$A_m\, h_{-2}\, t$	$A_c\, h_{-2}\, t$
Flicker FM	$h_{-1} f^{-1}$	$B\, h^{-1}$	$B_m\, h_{-1}$	$B_c\, h_{-1}$
White FM	h_0	$C\, h_0\, t^{-1}$	$C_m\, h_0\, t^{-1}$	$C_c\, h_0\, t^{-1}$
Flicker PM	$h_1 f$	$D\, h_1\, t^{-2}$	$D_m\, h_1\, t^{-2}$	$D_c\, h_1\, t^{-2}$
White PM	$h_2 f^2$	$E\, h_2\, t^{-2}$	$E_m\, h_2\, t^{-3}$	$E_c\, h_2\, t^{-3}$

TABLE 3.9

Coefficients to be used in Table 3.8

$A = \frac{2}{3}\pi^2$	$A_m = A_c$	$A_c = \frac{11}{20}\pi^2$
$B = 2\ln(2)$	$B_m = B_c$	$B_c = \frac{1}{8}(27\ln(3) - 32\ln(2))$
$C = \frac{1}{2}$	$C_m = C_c$	$C_c = \frac{1}{4}$
$D = \frac{1.038 + 3\ln(2\pi B\tau)}{4\pi^2}$	$D_m = \begin{cases} D & \text{small } n \\ D_c & \text{large } n \end{cases}$	$D_c = \frac{3}{8\pi^2}(8\ln(2) = 3\ln(3))$
$E = \frac{3B}{4\pi^2}$	$E_m = E\tau_0$	$E_c = \frac{1}{8\pi^2}$

If the modified Allan variance is measured without satisfied the sampling theorem, then aliasing occurs. The modified variance measured under this condition is noted “mod” in Table 3.8 and the coefficients A_m to E_m are to be used for the conversion (see Table 3.9). Contrary to all the other coefficients in the table these are not exact analytical values. The exact modified Allan variance depends on the sampling period τ_0 and on the number of samples n through equation (3.44). The coefficients shown are asymptotic values based on the assumption that for random walk FM, flicker FM and white FM the aliasing effect is negligible and that the discrete modified variance is equal to the continuous variance given by (3.46). For flicker PM it was verified in [Bernier, 1988] by numerical integration that the discrete modified variance behaves like the classical variance for small values of n and like the continuous modified variance for large values of n. In the case of white FM it was verified in [Bernier, 1987] by numerical integration that the discrete modified variance is equal to the classical Allan variance for $n = 1$ and follows the –3 slope of the continuous modified Allan variance even for small values of n. Conversion factors for non-integer exponent power-law processes are computed theoretically in [Walter, 1994] and fully confirm the coefficients presented here for integer exponents.

The process of translating from the frequency domain to the time domain is an exact one. However, the inverse process, going from the time domain to the frequency domain, is only approximate unless there is only one power-law process [Vernotte, 1993].

3.3.3 Environmental

The above measures can be used to characterise the environment and hence the effect of particular environmental parameters of an oscillator. In recent years the applications of precision oscillators have expanded very rapidly and the question of the characterisation of environmental influences is becoming increasingly important. The physics of environmental influences on frequency standards is treated in [Audoin et al, 1990; Mattison et al., 1976; Papoulis, 1983 and Walls, 1990]. The characterisation and the specification of environmental sensitivities is treated in [Beard et al., 1989; Breakiron, 1989; Brendel et al., 1989; Dragonette et al., 1991; Gagnepain, 1989; Garvey 1989; Sydnor, 1989; Walls, 1990 and IEEE, 1994].

3.4 A bridge to the next Chapter

Most of the formalism and notations used in this chapter are borrowed to the field of electrical engineering and telecommunication theory. The present chapter, therefore, does not follow the conventions normally used by the time and frequency specialists who have developed concepts and notations of their own. This section is a bridge from the notations used in the present chapter and the standard IEEE and ITU-R notations borrowed from the specialised time and frequency formalism that are used in the other chapters.

In the standard time and frequency notation, for example, the autocorrelation function of $\phi(t)$ is noted $R_\phi(\tau)$ and there is no distinction between the one sided and the two-sided power spectral densities. Both are noted $S_\phi(f)$.

Instead of (3.3) and (3.5), the standard model of the oscillator noise uses directly the projection of the phasor model on the real axis, i.e. the real signal instead of the analytic one:

$$u(t) = \mathrm{U}_0(1+\varepsilon(t))\sin(2\pi\nu_0 t + \phi(t)) \tag{3.65}$$

where $u(t)$ is the oscillator output voltage and U_0 the nominal amplitude.

3.5 Appendix: Random processes

3.5.1 Introduction

In the domain of frequency stability characterisation the quantities of interest are essentially band-limited, low-pass and band-pass random processes such as $x(t)$, $y(t)$ and $n(t)$ defined above.

3.5.2 Definition of a random process

A random process $x(t)$ is the statistical ensemble $\{\Omega\}$ of all the possible sample functions $x(\omega, t)$ that share the statistical properties that define the random process $x(t)$. The random process is to the sample function what the random variable is to a random number. A good introduction to random processes can be found in [Wozencraft et al., 1965].

The processes $\varepsilon(t)$, $\phi(t)$, $x(t)$, $y(t)$, $n(t)$ etc. defined in previous sections are all random processes.

3.5.3 Stationary random processes

A random process is stationary if its statistical properties are invariant to a translation of time and have a finite mean. This also means that its statistical properties are independent of the origin of time. A wide sense stationary process has a finite mean for statistical properties, but these are not invariant to a translation in time.

3.5.4 Non-stationary random processes

A random process is non-stationary if its statistical properties are a function of time. The phase noise $\phi(t)$ and the frequency noise $y(t)$ of oscillators generally contains non-stationary terms. For instance the phase noise process $\phi(t)$ usually has flicker or random walk components that makes its mean value and mean squared value undefinable. Only local averages, that are a random function of time, can be defined.

It is precisely the non-stationarity of the oscillator noise processes that ruled out the use of classical statistical methods for their characterisation in the time domain and that led to the development of specific methods such as the Allan variance. In all of the above, it should be kept in mind that stationarity or non-stationarity is a property of models and not of oscillators.

3.5.5 Auto-correlation function

The auto-correlation function of a real stationary random process $x(t)$ is defined as

$$R_{xx}(\tau) = E\{x(t)x(t+\tau)\} \tag{3.66}$$

where $E\{\}$ is the statistical expectation operator. In the case of a non-stationary or wide sense stationary real process the autocorrelation function is also a function of the time t and not only of the time difference τ.

3.5.6 Power spectral density

The double sided power spectral density of the process $x(t)$ is defined as the Fourier transform of its auto-correlation function.

$$S_{xx}(f)=\int_{-\infty}^{+\infty}R_{xx}(\tau)\exp(-j2\pi f\tau)d\tau=2\int_{0}^{+\infty}R_{xx}(\tau)\cos(2\pi f\tau)d\tau \tag{3.67}$$

The single sided power spectral density $S^{+}_{xx}(f)$ is the two-sided power spectral density folded over the origin so that only the positive Fourier frequencies are used.

$$S^{+}_{xx}(f)=\begin{matrix}2S_{xx}(f) & \text{for} & f\geq 0\\ 0 & \text{for} & f<0\end{matrix} \tag{3.68}$$

3.5.7 Linear filtering of random processes

By definition when a random process $z(t)$ is filtered by a linear operator of impulse response $h(t)$, the output $w(t)$ is the convolution of the impulse response $h(t)$ with the input $z(t)$. It can be shown that the power spectral density of the output $w(t)$ is given by [Wozencraft et al., 1965]

$$S^{+}_{ww}(f)=S^{+}_{xx}(f)\left|H(f)\right|^{2} \tag{3.69}$$

where $\left|H(f)\right|^{2}$ is the squared module of the transfer function $H(f)$ which is defined as the Fourier transform of the impulse response $h(t)$ of the linear operator.

References

ALLAN D.W. [1966] Statistics of Atomic Frequency Standards. *Proceedings IEEE*, Vol. 54, 221-230.

ALLAN D.W. [1987] Time and Frequency (Time-Domain) Characterization. Estimation, and Prediction of Precision Clocks and Oscillators. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. UFFC-34, no. 6, 647-654.

ALLAN D.W., BARNES J.A. [1981] A Modified 'Allan Variance' with Increased Oscillator Characterization Ability. Proceedings 35th Annual Frequency Control Symposium. Ft. Monmouth NJ, pp. 470-475.

ALLAN D.W., HELWIG H., KARTASCHOFF P., VANIER J., VIG J., WINKLER G.M.R., YANNONI N.F., [1988] Standard Terminology for Fundamental Frequency and Time Metrology. Proceedings 42nd Annual Symposium on Frequency Control, pp. 419-425.

ALLAN D.W., WEISS M.A., JESPERSEN J.L., [1991] A Frequency-Domain View of Time-Domain Characterization of Clocks and Time and Frequency Distribution Systems. Proceedings 45th Annual Symposium on Frequency Control, pp. 667-678.

AUDOIN C., DIMARCQ N., GIORDANO V., VIENNET J. [1990] Physical Origin of the Frequency Shifts in Cesium Beam Frequency Standards: related Environmental Sensitivities. Proceedings 22nd PTTI Meeting, Vienna (Virginia), pp. 419-440.

BARNES J.A., ALLAN D.W. [1989] Variances Based on Data with Dead-Time Between the Measurements: Theory and Tables. Technical Note 1318, National Bureau of Standards.

BEARD R., DANZY F., POWERS E., WHITE J. [1989] Test and Evaluation Methods at the NRL Clock Test Facility. Proceedings 43rd Annual Symposium on Frequency Control, pp. 275-288.

BEDROSIAN E. [1962] The Analytic Signal Representation of Modulated Waveforms. *Proceedings IRE*, Vol. 50, 2071-2076.

BERNIER L.G. [1987] Theoretical Analysis of the Modified Allan Variance. Proceedings 41st Annual Frequency Control Symposium, Philadelphia, pp. 116-121.

BERNIER L.G. [1988] Linear Prediction of the Non-Stationary Clock Error Function. Proceedings 2nd European Time and Frequency Forum, Neuchâtel CH, pp. 125-137.

BERNIER L.G., GARDIOL F.E. [1985] The Analytic Signal Representation of Oscillators with Application to Frequency Stability Analysis. Proceedings 39th Annual Frequency Control Symposium, Philadelphia, pp. 127-131.

BLACHMAN N.M., MCALPINE G.A. [1969] The Spectrum of a High-Index FM Waveform: Woodward's Theorem Revisited. *IEEE Transactions on Communication Technology*, Vol. COM-17, no. 2, 201-208.

BREAKIRON L. [1989] Measurement of the Frequency Response of Cesium Clocks to Temperature Fluctuations. Proceedings 43rd Annual Symposium on Frequency Control, pp. 296-297.

BRENDEL R., EL HASSANI C., BRUNET M., ROBERT E. [1989] Influence of Magnetic Field on Quartz Crystal Oscillators. Proceedings 43rd Annual Symposium on Frequency Control, pp. 268-274.

DRAGONETTE R.A., SUTER J.J. [1991] Barometric Pressure-Induced Frequency Offsets in Hydrogen Masers. Proceedings 45th Annual Symposium on Frequency Control, pp. 586-590.

GAGNEPAIN J. [1989] Characterization Methods for the Sensitivity of Quartz Oscillators to the Environment. Proceedings 43rd Annual Symposium on Frequency Control, pp. 242-247.

GARVEY R. [1989] Testing and Specification of Environmental Sensitivities in Cesium and Precision Quartz Signal Sources. Proceedings 43rd Annual Symposium on Frequency Control, pp. 263-267.

GREENHALL C.A. [1983] A Structure Function Representation Theorem with Applications to Frequency Stability Estimation. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-32, no. 2, 364-370.

GREENHALL C.A. [1989] A Method for Using a Time Interval Counter to Measure Frequency Stability. *IEEE Transactions. on Ultrasonics. Ferroelectrics and Frequency Control*, Vol. UFFC-36, no. 5, 478-480

GREENHALL C.A. [1992] A Shortcut for Computing the Modified Allan Variance. *IEEE Frequency Control Symposium,* pp. 262-264

IEEE [1988] Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology. IEEE Std 1139-1988.

IEEE [1994] Guide for Measurement of Environmental Sensitivities of Standard Frequency Generators. IEEE Std 1193-1994.

KARTASCHOFF P. [1987] Reference Clock Parameters for Digital Communications Systems Applications. Proceedings 19th PTTI Meeting, pp. 515-549.

LESAGE P. [1983] Characterization of Frequency Stability: Bias due to the Juxtaposition of Time-Interval Measurements. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-32., no. 1, 204-207.

LESAGE P., AUDOIN C. [1973] Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-22, 157-161.

LESAGE P., AUDOIN C. [1976] Corrections to: Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-25, 270.

LESAGE P., AUDOIN C. [1979] Effect of Dead Time on the Estimation of the Two-Sample Variance. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-28., no. 1, 6-10.

LESAGE P., AUDOIN C. [1979] Characterization and Measurement of Time and Frequency Stability. *Radio Science,* Vol. 14, no. 4, 521-539.

LINDSEY W.C., CHIE C.M. [1976] Theory of Oscillator Instability Based Upon Structure Functions. *IEEE Proceedings*, Vol. 64, 1662-1666.

MATTISON E.M. [1990] Physics of Systematic Frequency Variations in Hydrogen Masers. Proceedings 22nd PTTI Meeting, Vienna (Virginia), pp. 453-464.

NIST [1990] Characterization of Clocks and Oscillators. NIST Technical Note 1337.

PAGNELLI C.J., Cashin W.F. [1991] Measurement of Precision Oscillator Phase Noise Using the Two-Oscillator Coherent Down-Conversion Technique. Proceedings 23rd PTTI Meeting, Pasadena CA, NASA Conference Publication N° 3159, pp. 189-207.

PAPOULIS A. [1983] Random Modulation: A Review. IEEE Transactions Acoustics, Speech, and Signal Processing, Vol. ASSP-31, no. 1, pp. 96-105.

RILEY W.J. [1990] The Physics of the Environmental Sensitivity of Rubidium Gas Cell Atomic Frequency Standards. Proceedings 22nd PTTI Meeting, Vienna (Virginia), pp. 441-452.

RUTMAN J. [1978] Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen Years of Progress. *Proceedings IEEE*, Vol. 66, no. 9, 1048-1075.

STEIN S.R. [1985] Frequency and Time – Their Measurement and Characterization. *Precision Frequency Control*, Vol. 2, edited by E.A. Gerber and A. Ballato, Academic Press, New-York, 1985, pp. 191-416.

SYDNOR R. [1989] Environmental Testing at the Jet Propulsion Laboratory's Frequency Standards Laboratory. Proceedings 43rd Annual Symposium on Frequency Control, pp. 242-247.

TAVELLA P., PREMOLI [1991] Characterization of Frequency Standard Instability by Estimation of their Covariance Matrix. Proceedings PTTI Meeting.

VERNOTTE F., LANTZ E., GROSLAMBERT J., GAGNEPAIN J.J. [1993] Oscillator noise analysis: multi-variance measurement. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-42, no. 2, 342-350.

VERNOTTE F., MCHUGH M., ZALAMANSKY G. [1995] Cut-off frequencies and noise power law model of spectral density: adaptation of the multi-variance method using the structure function approach. Proceedings EFTF 95, Besançon, France, pp. 373-376.

WALLS F.L. [1990] Environmental Sensitivities of Quartz Crystal Oscillators. Proceedings 22nd PTTI Meeting, Vienna (Virginia), pp. 465-486.

WALLS F.L. and DEMARCHI A. [1975] RF Spectrum of a Signal After Frequency Multiplication; Measurement and Comparison with a Simple Calculation. *IEEE Transactions on Instrumentation and Measurement*, IM-24, pp. 210-217.

WALTER T. [1992] A multi-variance analysis in the time domain. Proceedings 24th PTTI.

WALTER T. [1994] Characterizing Frequency Stability: A Continuous Power-Law Model with Discrete Sampling. *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-43, no. 1, 69-79.

WOZENCRAFT J.M., JACOBS I.M. [1965] Principles of Communication Engineering, Wiley & Sons, New-York.

Chapter 4

Measurement techniques (Metrology)

Contents

Introduction

There are many measurement techniques for characterising the performance of precision frequency sources. These techniques differ widely in circuit design, the type of information available, uncertainty, and resolution. Table 4.1 adapted from [Sullivan et al., 1990] compares the approximate uncertainty and resolution for many of these methods for measuring time and frequency instabilities along with their disadvantages and advantages. Since Table 4.1 is very simple, it can not list all factors affecting the measurements and therefore should be considered only as a first step in selecting a measurement method. Additional details are given in following sections. These methods, grouped by measurement architecture are listed below.

To understand better the limitations for measuring time, frequency, and spectral purity it is necessary to discuss briefly definitions and measurement concepts. Much more lengthy discussions of the definitions and relation-ships between them are given in Chapter 3. The model for a signal voltage that incorporates the influence of noise is given in Eq. (4.1) [Sullivan et al., 1990, Barnes et al., 1971, Allan et al., 1988].

$$V(t) = (V_0 + \varepsilon(t))\cos(2\pi\nu_0 t + \phi(t)) \tag{4.1}$$

where V_0 is the average amplitude of the signal and ν_0 is the average frequency of the signal. Amplitude modulation (AM) noise is included in $\varepsilon(t)$ and phase modulation (PM) noise is included in the $\phi(t)$ term. Although instantaneous frequency (d/dt phase) is mathematically definable, in the real world it takes a finite amount of time to measure the phase slope. Therefore all frequency measurements yield a frequency averaged over a measurement time interval τ.

Equations (4.2) and (4.3) are working definitions for characterising the spectral purity of signals that are easily related to measurement techniques. Ordinarily one is most interested in $S_\phi(f)$, however, in many measurement configurations the noise floor or resolution is set by the presence of AM noise.

In many systems the PM performance is determined by AM to PM conversion. Therefore no determination of PM noise is complete without the estimation of the AM noise and its contribution to the apparent PM noise.

$$S_\phi(f) = \frac{(\delta\phi(f))^2}{BW} \tag{4.2}$$

where BW is the measurement bandwidth in Hz. $S_\phi(f)$ is the mean squared phase deviation separated by f from the carrier, measured in rad^2 per Hz measurement bandwidth.

TABLE 4.1
Guide to selection of measurement methods

Measurement Method	Time accuracy [a]	Time stability (1 Day)	Frequency accuracy [a]	Frequency stability [b] $\sigma_y(\tau)$	Advantages	Disadvantages
I. Direct measurements						
1. At the fundamental frequency	See Note [a]	Limited by time base stability	Limited by time base accuracy	Limited by time base stability	Very simple to perform	Extremely limited resolution not appropriate for high stability oscillators[d]
2. After multiplication/division	See Note [c]	Limited by time base stability	Limited by time base accuracy	Limited by time base stability	Very simple to perform; for frequency multiplication factor N, noise increases as 20 Log(N)	Provides only modest extension of above method and thus suffers similar limitations.[d]
II. Heterodyne measurements						
1. Single-conversion methods 2. Multiple conversion methods	-------	See Greenhall[e]	$\approx 10^{-16}$ at 10 MHz	$\approx 10^{-7}/(\nu_0\tau)$ which at 10 MHz is $10^{-14}/\tau$	Measurement Noise can typically be made less than oscillator instabilities for τ + 1 s and longer	Minimum τ determined by period of beat frequency, typically not adjustable; cannot compare oscillators near zero beat; additional information needed to tell which oscillator is high/ low in frequency; dead-time often associated with measurements
3. Time-difference method	~100 ps	~20 ps at 10 MHz	$\sim 10^{-16}$ is $10^{-14}/\tau$	$\approx 10^{-7}/(\nu_0\tau)$ which at 10 MHz is $10^{-14}/\tau$	Wide bandwidth input allows a variety of signals; simple to use; cycle ambiguity almost never a problem; measures time, time stability, frequency, and frequency stability	Using best available equipment, measurement noise is typically greater than oscillator instabilities for τ less than several seconds, hence it is often limited to long-term measurements.
3a. Dual-mixer time-difference method	~ 100 ps	~ 5 ps	$\sim 10^{-16}$ at 10 MHz	$\approx 10^{-7}/(\nu_0\tau)$ which at 10 MHz is $10^{-14}/\tau$	No dead-time; may choose sample time (1 ms to as large as desired); oscillators may be at zero beat or different; measurement bandwidth easily changed; measures time, time stability, frequency, and frequency stability.	More complex than other methods, and hence more susceptible to extraneous signal pick-up; e.g., ground loops; the time difference is modulo the beat period, e.g., 200 ns at 5 MHz.

[a] Accuracy of the measurement cannot be better than the stability of the measurement. Accuracy is limited by the accuracy of the reference oscillator.
[b] This is for a measurement bandwidth of 10^4 Hz; ν_o = frequency; τ = measurement time.
[c] This assumes use of a simple frequency counter.
[d] See [GREENHALL 1987]
The dash (-----) means that the method is not generally appropriate for this quantity.

TABLE 4.1

Guide to selection of measurement methods (continued)

Measurement method	Time accuracy [a]	Time stability (1 day)	Frequency accuracy [a]	Frequency stability [b] $\sigma_y(\tau)$	Advantages	Disadvantages
III. Homodyne methods					Particularly useful for measurement of phase noise	Generally not used for measurements of time
1. Phase-locked -loop methods					Assure continuous quadrature of signal and reference	Care needed to assure that noise of interest is outside the loop bandwidth
a. Loose-phase-locked-loop	________	________	Depends on calibration of varicap	$\sim 10^{-7}/(\nu_0\tau)$ which at 10 MHz is $\sim 10^{-14}/\tau$	Useful for short-term stability analysis as well as spectrum analysis and the detection of periodicity in noise as spectral lines; excellent sensitivity	Long-term phase measurements (beyond several seconds) are not practical
b. Tight-phase-locked-loop	________	________	Depends on calibration of varicap	$\sim 10^{-7}/(\nu_0\tau)$ which at 10 MHz is $\sim 10^{-14}/\tau$	Measurement noise typically less than oscillator instabilities for τ = 1 s and longer; good measurement system bandwidth control; dead-time can be made small or negligible	Need voltage controlled reference oscillator; frequency sensitivity is a function of varicap tuning curve, hence not conducive to measuring absolute frequency differences
2. Discriminator methods					Requires no reference oscillator	Substantially less bandwidth than two-oscillator, homodyne methods; sensitivity low at low Fourier frequency
a. Cavity discriminator	________	________	Depends on characteristics of discriminator	Depends on characteristics of discriminator	Requires no reference oscillator; very easy to set up and high sensitivity; practical at microwave frequencies	Requires more difficult calibration to obtain any accuracy over even modest range of Fourier frequencies; accurate only for Fourier frequencies less than 0.1 x bandwidth
b. Delay line	________	________	________	Depends on charateristics of delay line	Requires no reference oscillator; dynamic range set by properties of delay line; practical at microwave frequencies	Sunstantially less accurate than two-oscillator, homodyne methods; cumbersome sets of delay lines needed to cover much dynamic range; considerable delay needed for measurements below 100 kHz from carrier.

[a] Accuracy of the measurement cannot be better than the stability of the measurement. Accuracy is limited by the accuracy of the reference oscillator.

[b] This is for a measurement bandwidth of 10^4 frequency; τ = measurement time.

The dash (________) means that the method is not generally appropriate for this quantity.

Note #2 Appendix.

Table 4.2 lists the parameters which commonly affect the uncertainty of PM noise measurements.

TABLE 4.2

Error model for PM measurements [WALLS et al., 1988]

1.	Determination of k
2.	Determination of amplifier $G(f)$
3.	PLL effects (if any)
4.	Contribution of AM noise
5.	Harmonic distortion
6.	Contribution of system noise floor
7.	Contribution of reference noise
8.	Statistical confidence of data
9.	Linearity of spectrum analysers
10.	Uncertainty in PSD function

$$S_a(f) = \left(\frac{\delta\varepsilon(f)^2}{V_0} \right) \frac{1}{BW} \tag{4.3}$$

$S_a(f)$ is the fractional mean squared amplitude fluctuation per Hz bandwidth, separated from the carrier by frequency f. Table 4.3 lists the parameters which commonly affect the uncertainty of AM noise measurements.

TABLE 4.3

Error model for AM measurements [WALLS et al., 1988]

1.	Determination of k
2.	Determination of amplifier $G(f)$
3.	Contribution of system noise floor
4.	Statistical confidence of data
5.	Linearity of spectrum analysers
6.	Uncertainty in PSD function

Table 4.4 lists the confidence interval for spectral density measurements.

TABLE 4.4

Confidence interval for spectral density measurements

Statistical uncertainty of spectral density measurements as a function of $k\sqrt{\frac{\alpha}{N}}$, where k controls the confidence interval, α is the ratio of the video bandwidth to the resolution bandwidth for swept spectrum analysers and 1 for FFT spectrum analysers, N is the number of averages. To avoid biases the bandwidth must be very small compared to f. [WALLS et al., 1989; PERCIVAL et al., 1993; TAYLOR et al., 1993]

$\frac{N}{\alpha}$	$k = 1$ (approx. 68%) $S_m = S[1 \pm \delta], S_m \begin{smallmatrix} -\gamma \\ +\beta \end{smallmatrix}$ dB			$k = 1.9$ (approx. 90%) $S_m = S[1 \pm \delta], S_m \begin{smallmatrix} -\gamma \\ +\beta \end{smallmatrix}$ dB		
	δ	γ	β	δ	γ	β
4	0.54	–2	+3.3	2.5	–3	+6
6	0.42	–1.5	+2.3	1.4	–2.5	+5
10	0.32	–1.2	+1.7	0.61	–2.1	+4
30	0.18	–0.72	+0.86	0.35	–1.3	+1.8
100	0.1	–0.41	+0.46	0.19	–0.76	+0.92
200	0.058	–0.24	+0.25	0.14	–0.46	+0.51
1000	0.032	–0.13	+0.13	0.06	–0.26	+0.28
3000	0.018	–0.08	+0.08	0.035	–0.15	+0.15
10000	0.01	–0.04	+0.04	0.019	–0.08	+0.08

The fractional frequency stability is usually characterised by the Allan or two sample variance given by:

$$\sigma_y^2(\tau) = \frac{1}{2(N-2)\tau^2} \sum_{i=1}^{N-2} (x_{i+2} - 2x_{i+1} + x_i)^2 \tag{4.4a}$$

$$\sigma_y^2(\tau) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} (\bar{y}_{i+1} - \bar{y}_1)^2 \tag{4.4b}$$

$$\sigma_y^2(\tau) = \frac{2}{(\pi\nu_0\tau)^2} \int_0^\infty S_\varphi(f) \sin^4(\pi f\tau) \mathrm{d}f \tag{4.4c}$$

where N is the number of samples, x_i is the time deviation at the point i, M is the number of frequency deviation samples, y_i, averaged over τ and τ is the spacing between the time deviation measurements [Barnes et al., 1971; Allan et al., 1988]. Equation (4.4a) is used for timing data, (4.4b) is used for frequency data, and (4.4c) is used for PM noise data. If the dominant noise type in short-term is flicker phase modulation (PM) or white PM, the modified Allan variance given by Eqs. (3.39), (3.40), or (3.44) of Chapter 3 can be used to improve the estimate of the underlying frequency stability of the sources [Barnes et al., 1971; Allan et al., 1988; CCIR, 1986; Walls et al., 1975; Stein, 1985; Rutman et al., 1991; Allan et al., 1981; Lesage et al., 1984; Walls et al., 1990; Bernier, 1987; Weiss, 1995]. See references [Stein, 1985; Rutman et al., 1991] for further discussion.

Figure 4.1 shows the shape of $\sigma_y(\tau)$ for various power-law noise types.

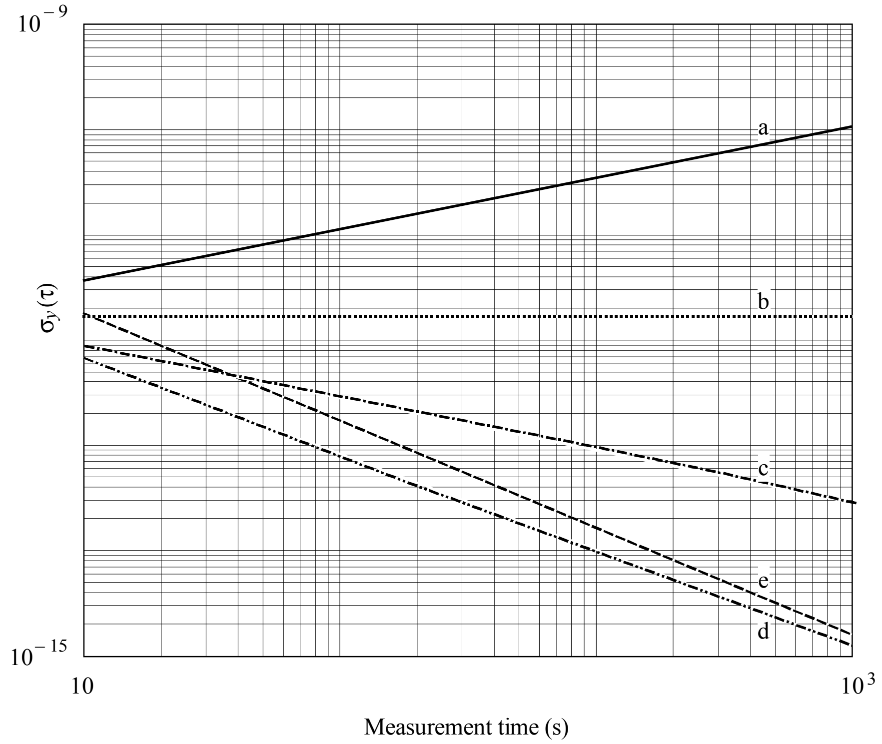


FIGURE 4.1

$\sigma_y(\tau)$ versus (τ) for the five common power-law noise types

Figure 4.2 shows the ratio of mod $\sigma_y^2(n\tau_0)$ to $\sigma_y^2(n\tau_0)$ for the 5 common noise types, in the limit that $2\pi f_h\,\tau_0 >> 1$.

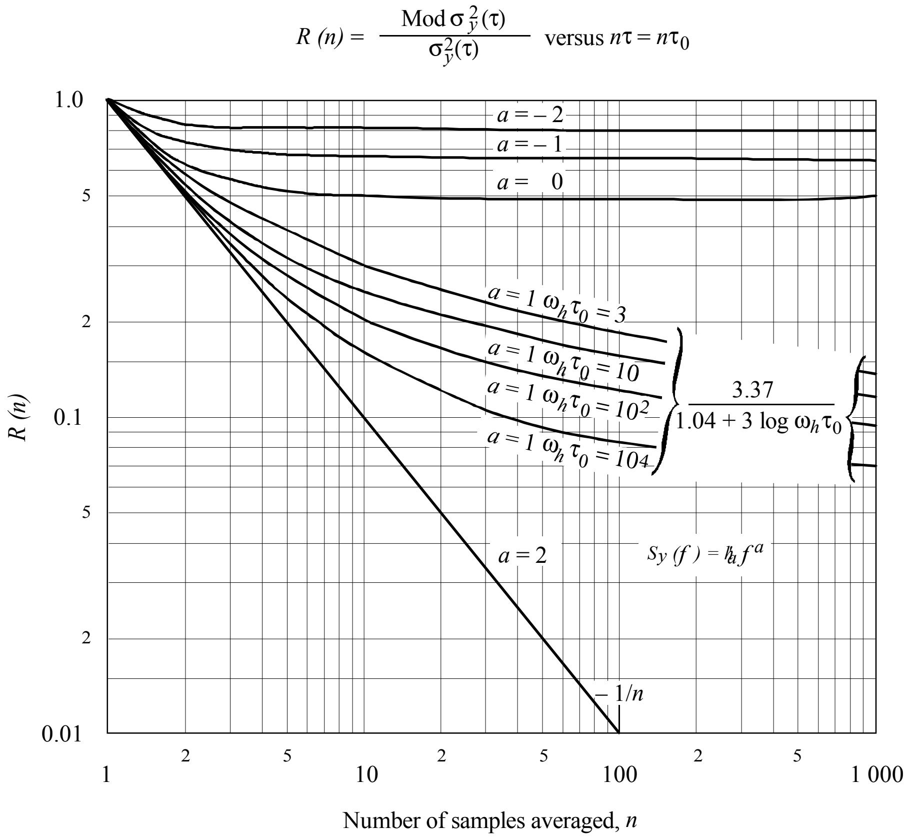


FIGURE 4.2

Ratio of mod $\sigma_y^2(n\tau_0)$ to $\sigma_y^2(n\tau_0)$ as a function of n

Tables 4.5 lists the parameters which commonly affect the uncertainty in $\sigma_y(n\tau)$ and mod $\sigma_y(\tau)$ measurements.

TABLE 4.5

Error model for $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$ measurements

1	Dead time
2	Measurement bandwidth
3	Contribution of refernece source or time base
4	Noise in zero-crossing detector (time-domain)
5	Noise in PM measurement system (frequency-domain)
6	Frequency drift in sources
7	Environmental effects, e.g., changes in mixer offset due to changes in temperature, humidity, and load
8	Phase-pulling from other sources

Table 4.6 lists the 68% confidence intervals for $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$ for full overlap of the data. See [Walter, 1994 and Weiss et al., 1993] for further details. [Howe, 1995] has developed a new method for analysing data to calculate $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$ which eliminates the bias introduced by removing an average frequency offset from the data. This greatly improves the confidence interval when the measurement time is longer than 20% of the data set.

TABLE 4.6

One sigma (68%) confidence intervals for time domain measurements

One sigma (68%) confidence intervals for time domain measurements with 1025 samples spaced τ_0 apart. Measurement times are given at $\tau = m\tau_0$. Columns 2 and 5 give the approximate degrees of freedom for $\sigma_y(\tau)$ and mod $\sigma_y(\tau)$ respectively [Stein, 1985; Walters, 1995, and Weiss, et al., 95]. Columns 3 and 4 give the one sigma confidence intervals for $\sigma_y(\tau)$ with full overlap of the data calculated from [Stein, 1985 and Howe et al., 1981]. Columns 6 and 7 give the one sigma confidence intervals for mod $\sigma_y(\tau)$ with full overlap of the data [WEISS et al., 95]. The work of [Howe, 1995] removes a long term bias and improves the confidence intervals when the measurement time is longer than 20% of the total data length.

n=1025	$\sigma_y(\tau)$ Deg. of Freed. White PM	Full Overlap – for 68% $\sigma_y(\tau)$ White PM	Full Overlap + for 68% $\sigma_y(\tau)$ White PM	mod $\sigma_y(\tau)$ Deg. Of Freed. White PM	Full Overlap – for 68% mod $\sigma_y(\tau)$ White PM	Full Overlap + for 68% mod $\sigma_y(\tau)$ White PM
m=1	526	2.9%	3.2%	526	2.9%	3.2%
m=2	526	2.9%	3.2%	477	3.1%	3.4%
m=4	524	2.9%	3.2%	299	3.8%	4.3%
m=8	521	2.9%	3.2%	158	5.2%	6.1%
m=16	515	3.0%	3.3%	78.9	7.1%	9.0%
m=32	503	3.0%	3.4%	38.2	9.7%	14%
m=64	479	3.0%	3.5%	17.6	13.%	22%
m=128	432	3.1%	3.6%	7.40	18.%	41%
m=256	355	3.4%	4.0%	2.85	24.%	94%
n=1025	Flicker PM	Flicker PM	Flicker PM	Flicker PM	Flicker PM	Flicker PM
m=1	590	2.8%	3.0%	590	2.8%	3.0%
m=2	554	2.9%	3.1%	497	3.0%	3.3%
m=4	453	3.2%	3.5%	263	4.1%	4.6%
m=8	336	3.6%	4.0%	128	5.7%	6.8%
m=16	232	4.3%	5.0%	62.3	7.8%	10%
m=32	151	5.2%	6.1%	29.8	11%	16%
m=64	92.3	6.7%	8.4%	13.7	15%	26%
m=128	52.1	8.4%	11%	5.74	20%	50%
m=256	26.2	11%	16%	2.07	26%	134.%
n=1025	White FM	White FM	White FM	White FM	White FM	White FM
m=1	682	2.6%	2.8%	682	2.6%	2.8%
m=2	584	2.8%	3.0%	516	3.0%	3.2%
m=4	354	3.5%	4.0%	252	4.1%	4.7%
m=8	186	4.8%	5.6%	123	5.8%	7.0%
m=16	93.5	6.4%	8.1%	59.8	8.0%	10%
m=32	45.9	8.8%	12%	28.7	11%	16%
m=64	22.0	12%	19%	13.2	15%	27%
m=128	10.0	16%	32%	5.50	20%	51%
m=256	4.0	22%	65%	1.81	27%	158%
n=1025	Flicker FM	Flicker FM	Flicker FM	Flicker FM	Flicker FM	Flicker FM
m=1	829	2.4%	2.5%	829	2.4%	2.5%
m=2	606	2.6%	3.0%	524	2.9%	3.2%
m=4	307	3.8%	4.3%	246	4.2%	4.8%
m=8	150	5.1%	6.0%	120	5.8%	7.1%
m=16	73.5	7.1%	9.0%	58.5	8.0%	11%
m=32	35.8	9.9%	14%	28.0	11%	16%
m=64	17.0	13%	22%	12.9	15%	27%
m=128	7.62	18%	41%	5.31	20%	53%
m=256	3.01	24%	90%	1.56	27%	192%
n=1025	Rand. Walk FM	Rand. Walk FM	Rand. Walk FM	Rand. Walk FM	Rand. Walk FM	Rand. Walk FM
m=1	1023	2.1%	2.2%	1023	2.1%	2.2%
m=2	511	3.0%	3.3%	442	3.2%	3.5%
m=4	254	4.1%	4.8%	200	4.6%	5.4%
m=8	125	5.7%	7.0%	97.2	6.4%	8.0%
m=16	61.2	7.8%	10%	47.3	8.8%	12%
m=32	29.2	11%	16%	22.6	12%	19%
m=64	13.3	15%	26%	10.3	16%	32%
m=128	5.51	20%	51%	4.19	22%	65%
m=256	2.0	26%	134%	1.29	28%	256%

4.1 Direct measurements of time (phase) and frequency

This technique is characterised by a direct comparison of the phase of the signal under test to that of a reference without the use of mixers. The timing diagram of the measurement is shown in Figure 4.3a.

4.1.1 Direct measurements of time (phase)

Time uncertainty and stability are fundamentally limited by the accuracy and stability of the reference signal and technically limited by the time resolution of the counter. Typical resolutions for time interval counters vary from picosecond to microsecond. Measuring time requires great care to assure that the voltage-standing-wave-ratio (VSWR) is small so that the phase of both the signal and reference is meaningful and reproducible [Nelson et al., 1992]. For sine wave signals, time is usually referenced to the positive going zero crossing of the signal. (For digital signals, time is usually referenced to the mean of the "0" and "1" states.) The counter may be started with the signal being measured or with the reference. Logically, we usually start with the signal then advancing time (or phase) corresponds to a frequency which is higher than the reference. The resolution is limited to $1/n$ where n is the frequency of the counter time-base oscillator. See Figure 4.3a. In some cases the reference may provide this time base as well as the stop signal, for example. Interpolation techniques in more sophisticated counters can increase the resolution by as much as a factor of 100.

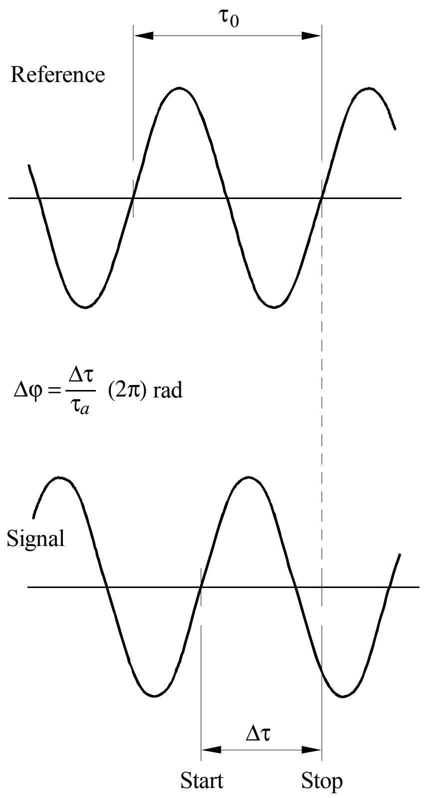


FIGURE 4.3a

Timing diagram of direct time (phase) measurement system

The time-domain fractional frequency stability of the reference and signal pair can be estimated from the timing data by using Eq. (4.4a). Since the short-term resolution of this approach is usually limited by the counter resolution, the use of the mod $\sigma_y(\tau)$ (Eq. 4.39) Chapter 3 is often helpful in obtaining a better estimate of the underlying clock instabilities. These measurements usually have dead time between measurements which causes a bias in the estimates of fractional frequency stability [Barnes et al., 1990]. See

Chapter 3 for a discussion. Figure 4.3b illustrates a straight-forward time interval measurement between two one pulse per second (1 PPS) signals. It is important to appreciate that 1 PPS signals are derived from the oscillator signal (5 MHz, for example) by some kind of counting and/or dividing circuit. For example, after 5 million positive-going zero crossings occur from a 5 MHz oscillator a circuit produces a pulse. This circuit often degrades the time stability. In addition, the derived 1 PPS signal usually has a very fast rise time. This requires an appropriate high-frequency front end for the time-interval counter.

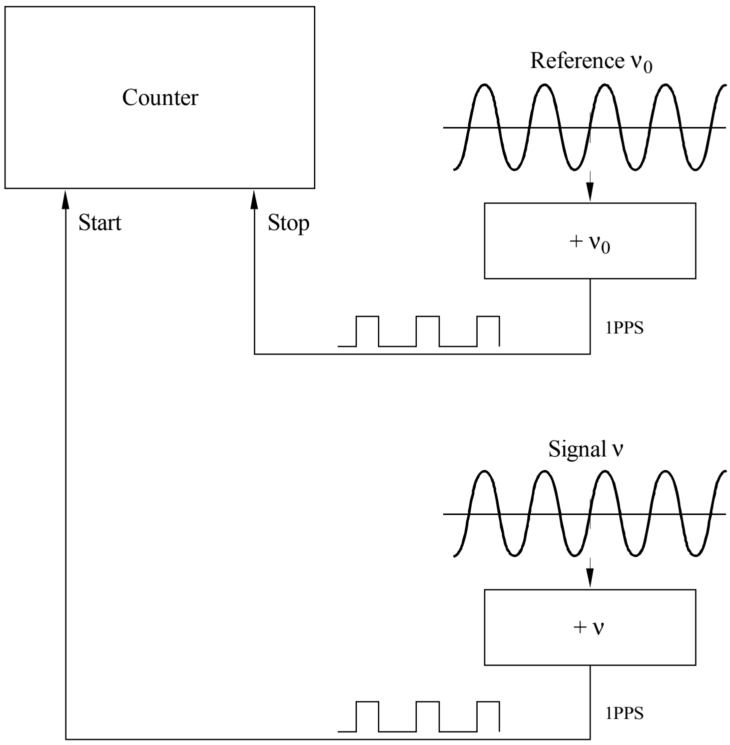


FIGURE 4.3b

Timing diagram of 1 pulse-per-second systems

4.1.2 Direct measurements of frequency

In the simplest implementation of this approach, the reference is used as the time base of a counter. The instrument reports the number of full cycles that occur during the specified number of cycles (zero crossings) of the reference as illustrated in Figure 4.4. The instrument usually counts the number of zero crossings of the measured signal for a given gate time reference. This measurement is limited in resolution to one cycle of the reference. For example, if the reference were 10 MHz, the normalised resolution of the time base would be $1/(10\ \text{MHz} \cdot \tau)$, where τ is the counters gate time. The measurement is also limited to one cycle of the measured signal. In more complex counters, interpolation techniques are used to estimate the number of full cycles plus the fractions of a cycle. For frequencies smaller than the reference frequency, the signal is sometimes used as the time base of the counter counting the reference. The results are then inverted to obtain the frequency of the signal under test. The uncertainty for long measurement times using either approach is limited by the relative standard combined uncertainty of the reference inaccuracy. Short-term resolution is limited by both the stability of the reference and the period of the reference or signal.

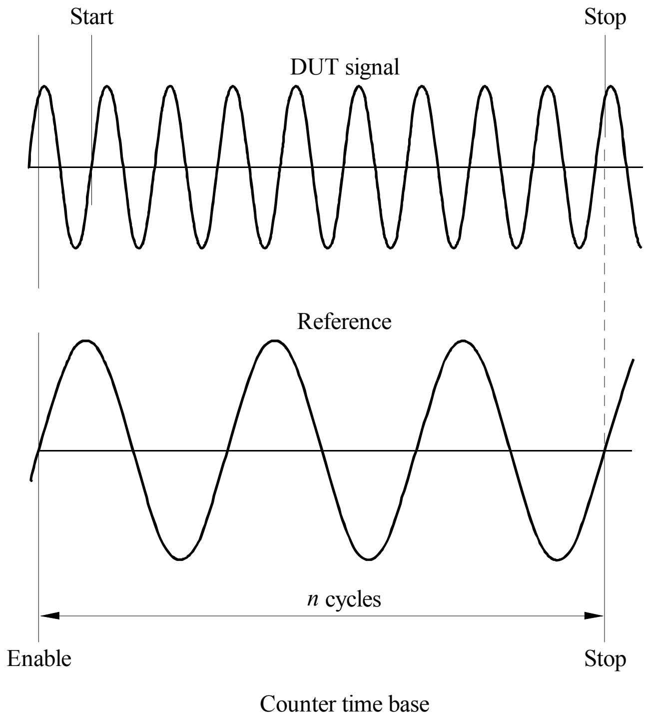


FIGURE 4.4

Timing diagram of direct frequency measurement system

The time-domain fractional frequency stability of the signal and reference can be estimated from the frequency data using Eq. (4.4b). Since the short-term resolution of this approach is usually limited by the counter resolution, the use of Eq. (3.40) of Chapter 3 is often helpful in obtaining a better estimate of the underlying instability of the sources. These measurements usually have dead time between measurements. The bias of the results due to dead time depends on noise type [Barnes et al., 1990].

4.2 Heterodyne measurements of frequency and phase (time)

Heterodyne techniques offer greatly improved short-term resolution over direct measurement techniques of time and frequency and are often used to measure the PM noise or frequency domain representation of frequency stability. In the heterodyne technique the signal under test ν is heterodyned against the reference signal ν_o and the difference or beat signal ν_b is measured. The various signals are shown in Figure 4.5. The frequency resolution is improved by a factor ν_o/ν_b over direct measurements, where $\nu_b = \nu_{signal} - \nu_0$.

The beat signal is usually obtained by connecting the two signals to a non-linear element such as a double balanced mixer. The output of the mixer yields the sum and difference frequency plus harmonics. The difference is extracted by low pass filtering and measured. The noise introduced by the mixer can in principle limit the short-term resolution but is usually overshadowed by the noise in available sources. When PM noise is being measured with analogue techniques, the beat frequency is usually set to zero. When digital techniques are used, the beat frequency is adjusted to be at least twice as high as the highest Fourier frequency of interest for PM or AM noise. The digital approach has the advantage that PM, AM, and the frequency stability can all be measured with the same set-up. Resolution, however, is usually not as good as the best analogue approaches.

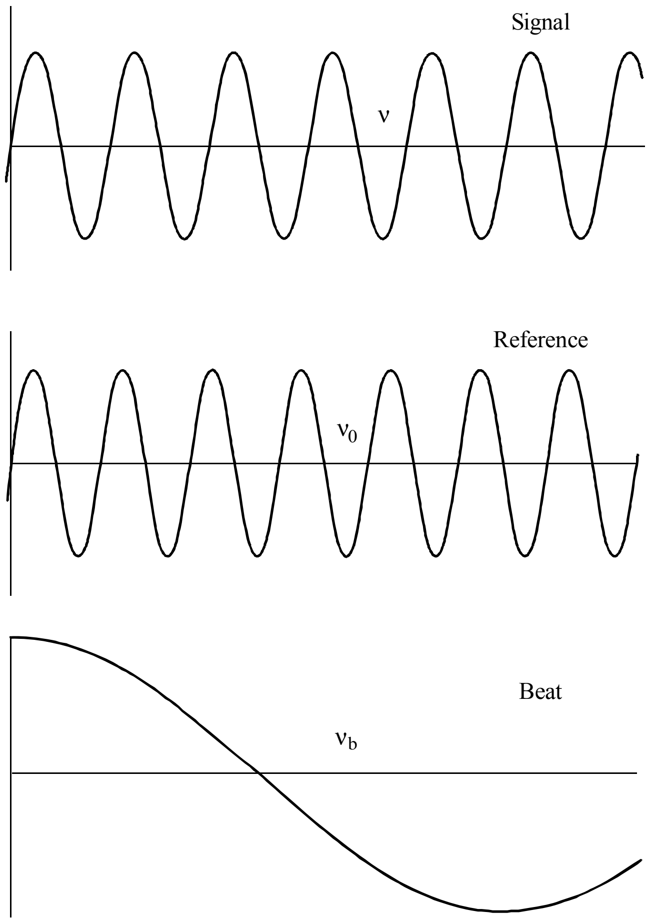


FIGURE 4.5

Timing diagram of hetorodyne time measurement system

The beat signal is usually obtained by connecting the two signals to a non-linear element such as a double balanced mixer. The output of the mixer yields the sum and difference frequency plus harmonics. The difference is extracted by low pass filtering and measured. The noise introduced by the mixer can in principle limit the short-term resolution but is usually overshadowed by the noise in available sources. When PM noise is being measured with analogue techniques, the beat frequency is usually set to zero. When digital techniques are used, the beat frequency is adjusted to be at least twice as high as the highest Fourier frequency of interest for PM or AM noise. The digital approach has the advantage that PM, AM, and the frequency stability can all be measured with the same set-up. Resolution, however, is usually not as good as the best analogue approaches.

Care must be exercised in heterodyne measurements that neither source is perturbed by the phase of the other one. The problem is usually approached by using a high isolation distribution amplifier on the output of each source. The maximum perturbations for an isolation γ in dB is:

$$\delta\phi_{max} = 10^{-\gamma/20} \quad \text{rad} \tag{4.5}$$

If the frequency offset between the sources is less than $\phi_{max}/2Q_1$, where Q_1 is the loaded Q-factor of the source, the phases will track "injection lock," yielding virtually no accumulated phase advance for extended times. It is a good practice to purposefully have the clock output frequencies be different so that if injection locking is a significant problem, the phase accumulation will appear as a series of discreet values followed by a jump in the value.

A significant advantage of a two oscillator measurement system is that the noise floor can be verified using an inferior oscillator as explained in Section 4.2.3 [Walls et al., 1988].

4.2.1 Heterodyne measurements of phase (time)

The resolution for heterodyne measurements of time or phase is increased to:

$$\delta\tau = (\nu_0 / \nu_b)\,\delta t \tag{4.6}$$

where δt is the resolution of the counter as discussed in Section 4.1.1. To avoid ambiguity ν_b should be compared to the peak frequency variations between the reference and the source under test. Additional measurements are required to determine if the frequency of the source is higher or lower than the reference. The phase of the beat signal goes through zero when the phase difference between the two signals is $\pm(2n + 1)$ 90º where $n = 0, 1, 2, 3,...$ See Figure 4.5. The time of the zero crossing is biased or in error by $\delta\phi$ due to imperfections in the symmetry of the mixer and/or the VSWR in the reference and signal paths [Nelson et al., 1992]. Timing errors due to VSWR effects and typical temperature coefficients for mixer biases are given in [Nelson et al., 1992] for frequencies of 5 and 100 MHz. These errors generally scale as $1/\nu$.

The time of the source under test is:

$$t_{DUT} = T_{ref} \pm n\nu_0 \pm \delta\phi \tag{4.7}$$

where n is the number of beat cycles that have occurred since the original synchronisation. The minimum time between data samples is $1/\nu_b$. For clocks that are at nearly the same frequency, this limitation can be very restrictive.

The time difference data can be used to characterise the fractional frequency stability of the sources using Eq. (4.4a) or (Eq. 3.39) of Chapter 3. The resolution for short-term time-domain frequency stability (τ less than 0.1 s) is typically much worse than that obtained from integrating the phase noise using Eq. (4.5c) [Walls et al., 1990].

4.2.2 Heterodyne measurements of frequency

Using the heterodyne method, the frequency of the DUT is:

$$\nu_{DUT} = \nu_0 \pm \nu_b \tag{4.8}$$

Additional measurements are required to determine the sign of the frequency difference. One method is to change the frequency of the reference or the DUT by a known amount and determine if the beat becomes smaller or larger. Another way is to add or subtract a phase (or time) shift to the reference or the DUT and observe the direction of the shift in the beat frequency. The resolution for a heterodyne frequency measurement is given by:

$$\delta\nu = \delta t \frac{\nu_b^2}{\nu_0} \tag{4.9}$$

where δt is the timing resolution for $\tau = 1/\nu_b$. See Section 4.1.1 for a discussion. Typically, the uncertainty is limited by the frequency stability of the reference, the phase variations of the phase detector, and other factors listed in Tables 4.2 and 4.4.

After a count is finished, a typical counter will arm and wait for the next zero-crossing to begin the next count. Hence in a slow beat-frequency situation, half the time the counter is not taking data (waiting for the next zero-crossing). This down time, referred to as “dead-time” [Barnes et al., 1990], biases the calculation of $\sigma_y(t)$ and mod $\sigma_y(t)$ by an amount that depends on noise type and length of the dead-time. Bias tables as a function of noise type and percent dead-time are given in [Barnes et al., 1990]. The dead-time limitation can be circumvented by using two counters triggered on alternate cycles or by using one of the techniques discussed in Sections 4.2.4-4.2.6. For clocks that have nearly the same frequency this is a serious limitation.

4.2.3 Heterodyne measurements of PM noise

Figure 4.6 shows the block diagram of a typical analogue measurement system used to measure PM noise of a source versus a reference oscillator. The mixer is used as a phase detector to transform small phase variations into small voltage fluctuations which are then measured by a spectrum analyser. Tables 4.2 and 4.4 show the parameter which generally affect measurement uncertainty. Figure 4.5 shows the typical output of the mixer versus phase difference between the source and the reference. The sensitivity of the mixer k_d and nominal amplifier gain $G(f)$, assuming the measurements are made after the amplifier, is determined from Eq. (4.10), the measurement of the slope of the beat signal at the zero crossing, and the beat period.

$$k_d G(f) = \left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{v=0} \frac{Tb}{2\pi} \tag{4.10}$$

To obtain a linear transfer function from the mixer it is generally necessary to maintain the phase fluctuations near 90° (quadrature) where the output voltage is approximately zero. Maintaining the phase deviations from quadrature smaller than 0.1 radian generally reduces the error below 0.2 dB. This is usually accomplished by using a phase-locked-loop (PLL) which reduces the phase fluctuations at low Fourier frequencies and passes Fourier frequencies higher than the bandwidth of the PLL [Walls et al., 1988; Howe et al., 1981; Walls et al., 1976]. Usually a second order PLL (one with an integrator) is used to minimise the phase error [Walls et al., 1976]. See Figure 4.6.

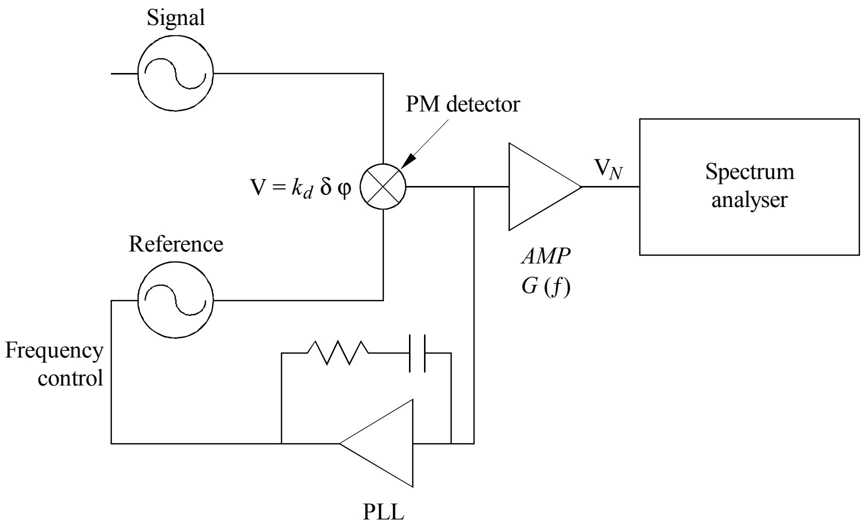


FIGURE 4.6

Block diagram for measurements of PM noise measurement system using a phase-locked-loop

Important errors arise from the frequency dependence of the mixer and the amplifier following the mixer [Walls et al., 1988]. The noise of the mixer and post amplifier set the resolution or noise floor of this configuration. Reactive terminations improve the mixer sensitivity thereby improving the noise floor, however the frequency response is not as flat and the phase errors are increased over that obtained with 50 ohm terminations [Nelson et al., 1992; Walls et al., 1988]. Figure 4.7 shows typical dependence of a low level mixer at 5 MHz on power and capacitive loading.

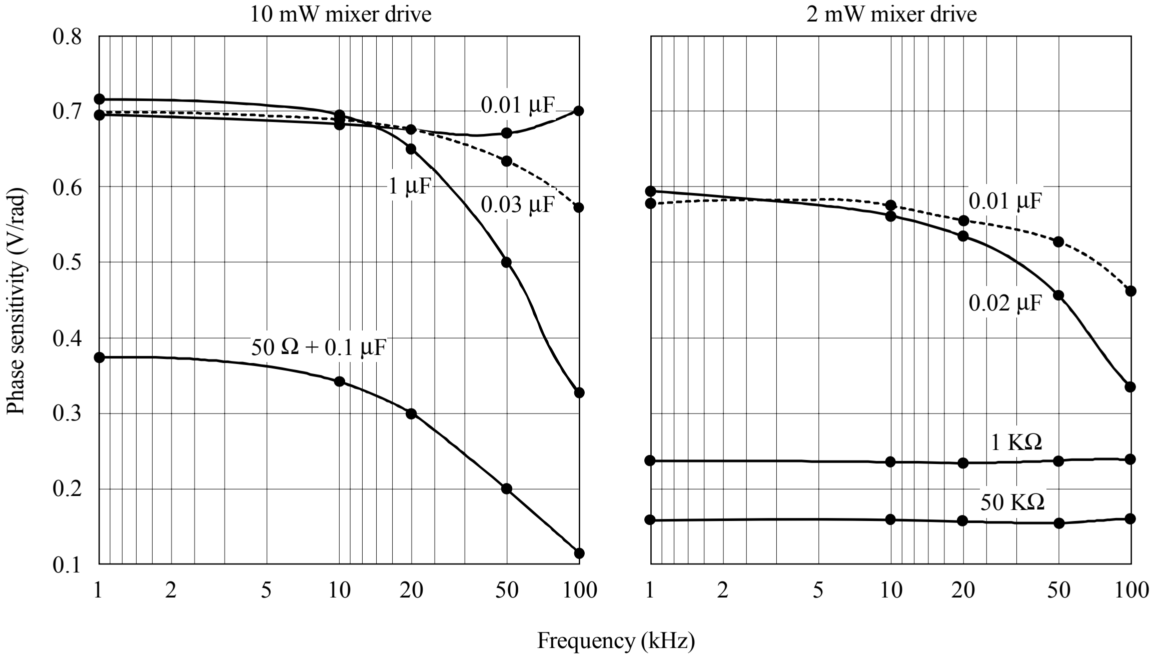


FIGURE 4.7

Typical sensitivity of a low level double balanced mixer at 5 MHz as a function of IF termination for drive levels of +2 and +10 dBm

The errors and uncertainties due to these effects can be significantly reduced by using one of the following calibration techniques. A very versatile technique is shown in Figure 4.8 [Walls et al., 1988; Walls, 1992; Walls et al., 1991].

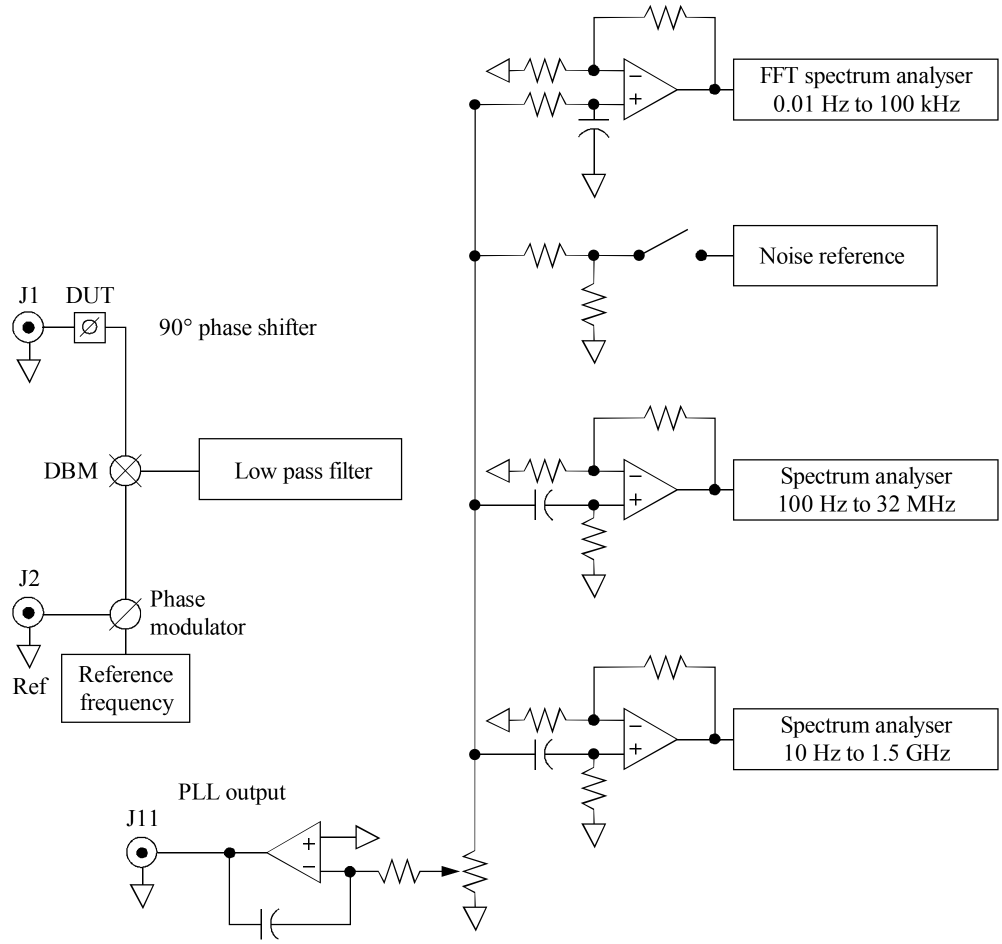


FIGURE 4.8

Block diagram of NIST extended PM measurement system using internal phase modulator to calibrate mixer sensitivity and amplifier sensitivity versus *f*

In this technique phase modulation sidebands, which are of approximately constant amplitude with f, are encoded on the signal by the modulator. The frequency of the modulation source is swept over the Fourier offset frequencies of interest. Since the amplitude of the phase modulation sidebands is constant versus Fourier offset frequency, the amplitude of the demodulated signal at the spectrum analyser can be used to correct for all the frequency dependent errors [Walls et al., 1988]. The mixer sensitivity and gain at low frequencies is determined at the beat frequency by measuring the slope of the beat signal after the amplifier in Figure 4.8. The approach of Figure 4.8 greatly improves accuracy for wideband measurement systems. Measurements with a standard certainty of less than 2 dB have been reported for Fourier frequencies of up to 1 GHz from the carrier [Walls et al., 1988; Walls, 1992; Walls et al., 1991].

A new technique for calibrating the frequency dependent errors that is well adapted to systems where many measurements are made at the same frequency, is shown in Figure 4.9. In this technique, a small amount of Gaussian noise centred about the carrier is added to the reference in a low noise power summer. This results in an equal amount of PM and AM noise which can be added at will to the reference [Walls, 1993-1; Walls, 1993-2; Walls et al., 1994]. When the added noise is off, the degradation of the reference PM noise is virtually unmeasureable [Walls, 1993-2]. The source to be measured is phase locked to the reference using a PLL as described above. No beat frequency data is required when using this technique.

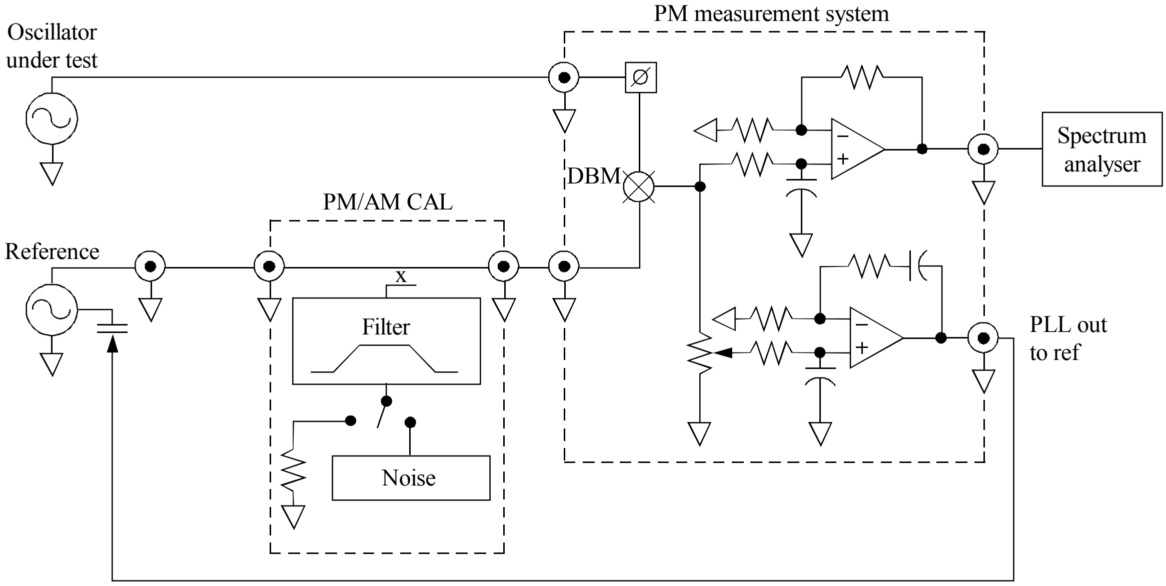


FIGURE 4.9

Block diagram showing the use of NIST PM/AM CAL noise standard to determine the PM noise of a oscilator

The added PM/AM noise can be made exceptionally constant with Fourier frequency as is shown in Table 4.7. The PSD of the noise voltage V_n is measured with noise on and noise off. The PSD with noise on is equal to the calibrated noise level multiplied by the gain of the mixer and amplifiers. We assume here that the noise of the oscillators and the system noise can be neglected in this measurement and that the phase detector has sufficient rejection of the AM noise so that its contribution can be neglected. Measurements where the system noise is significant are discussed in Section 4.2.8. The PSD of V_n with the noise off is equal to the noise of the oscillators under test plus the system noise multiplied by the power gain of the mixer amplifier system. If the PSDs are expressed in dB,

$$S_\phi(f)_{DUT} = S_\phi(f)_{calib} - \text{the difference} \tag{4.11}$$

as illustrated in Figure 4.10. The variations in PM/AM CAL reflect changes in overall sensitivity due to frequency dependent effects such as the action of the PLL and amplifier gain. $L(f)$ of the device under test (DUT) is obtained from eq. (4.11) [Walls, 1990; Walls et al., 1991; Walls, 1993-1; Walls, 1993-2]. This approach greatly reduces the uncertainty of the measurement because it automatically takes into account all the error terms, even the frequency dependent ones are accounted for, except AM to PM conversion. This approach also reduces the amount of time necessary to make routine PM noise measurements as compared to traditional methods since the measurement is now reduced to a ratio measurement between noise on and noise off [Walls, 1993-2].

TABLE 4.7

Noise Characteristics for 5 MHz, 10 MHz, 100 MHz, and 10.6 GHz PM/AM Noise Standards

Nominal source Phase Noise/Channel, ± 3dBc/Hz								
Fourier Frequency								
Source Frequency	1 Hz	10 Hz	100 Hz	1 kHz	10 kHz	100 kHz	1 MHz	10 MHz
5 MHz	–121	–151	–163	–171	–174	–174	–174	
10 MHz	–115	–145	–157	–165	–168	–168	–168	
100 MHz	–70	–100	–130	–156	–170	–170	–173	–173
10.6 GHz	+30	0	–30	–60	–85	–110	–140	–169
Maximum Residual Noise Between Channels, dBc/Hz								
5 MHz	–162	–172	–182	–190	–194	≤–175	≤–175	
10 MHz	–161	–176	–183	–191	–197	≤–175	≤–175	
100 MHz	–152	–162	–172	–182	–193	–193	–194	
10.6 GHz		–153	–163	–173	–181	–181	–196	–198
Differential PM/AM Noise, ±0.2 dBc/Hz								
5 MHz	–127.3	–127.3	–127.3	–127.3	–127.3	–127.3		
10 MHz	–128.4	–128.4	–128.4	–128.4	–128.4	–128.4	–128.4	
100 MHz	–129.5	–129.5	–129.5	–129.5	–129.5	–129.5	–129.5	–129.8
10.6 GHz	–138.9	–138.9	–138.9	–138.9	–138.9	–138.9	–138.9	–138.9

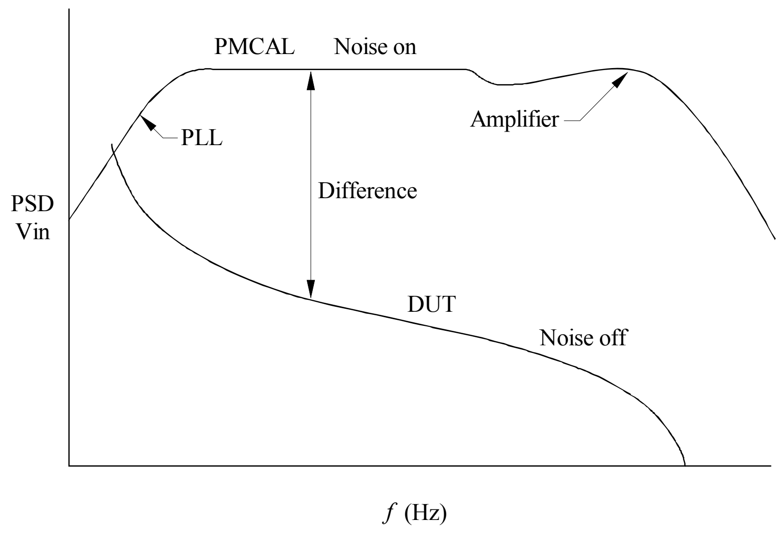


FIGURE 4.10

Calibration determination using PM/AM CAL from Figure 4.9

The noise floor of two oscillator measurement systems can be determined by using the scheme shown in Figure 4.11. A reactive power splitter is used to provide a reference signal for both ports of the PM measurement system. Since the phase between the two signals is now fixed, a third signal is needed to calibrate the mixer sensitivity and the amplifier gain. One approach is to set the phase of ϕ so that the mixer output is nearly zero.

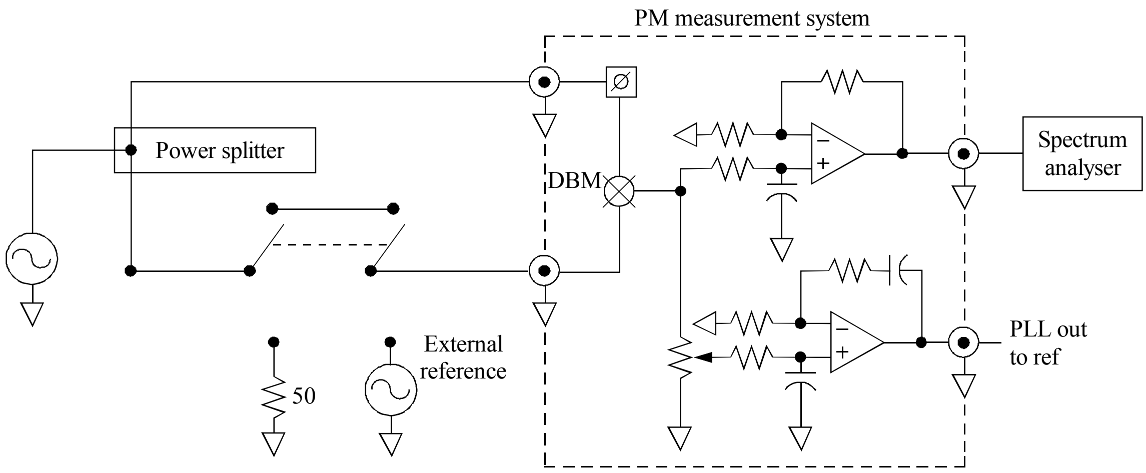


FIGURE 4.11

Block diagram of conventional configuration to measure the noise floor of PM noise measurement systems

One port of the power splitter is then terminated with 50 ohms and the cable attached to another source with the same output impedance and power as the output of the power splitter. The external source is offset from the reference to obtain a convenient beat signal to calibrate the mixer sensitivity and amplifier gain as per Eq. 4.10. After the calibration of k_d $G(f)$, the cable is reattached to original reference and the noise floor determined from:

$$\frac{PSD\ V_N}{[k_d G(f)]^2} = S_{\phi meas.syst.} + 2\beta^2 S_{a1}(f) + \frac{\pi f}{\nu_0} S_{\phi 1}(f) \tag{4.12}$$

where $S_{\phi 1}(f)$ is the PM noise of the reference, β^2 is the AM to PM conversion of the mixer, and $S_{a1}(f)$ is the AM noise of the reference. We have assumed the minimum phase shift necessary to obtain phase quadrature of the mixer.

A major limitation of this approach for wideband measurement systems is that the calibration does not measure the frequency variations of k_d or $G(f)$ or the contribution of the reference noise [Walls et al., 1988]. See Sections 4.2.7 and 4.2.8 for a discussion.

A new technique which generally has higher accuracy, because no substitution source or beat frequency measurements are needed and the frequency dependence for k_d and $G(f)$ are automatically included, is shown in Figure 4.12. In this approach the mixer sensitivity and amplifier gain are calibrated by measuring $PSD\ V_n$ with the PM/AM CAL noise standard on. The noise source is turned off and the $PSD\ V_n$ due to the residual noise is measured. The noise floor is given by Eq. (4.11). See Figure 4.10. The noise floor of this approach has been shown to be extremely low approaching – 195 dBc/Hz in some cases. Table 4.7 shows an example of the PM noise floors achieved using this technique for carrier frequency from 5 MHz to 10.6 GHz [Walls, 1993-2].

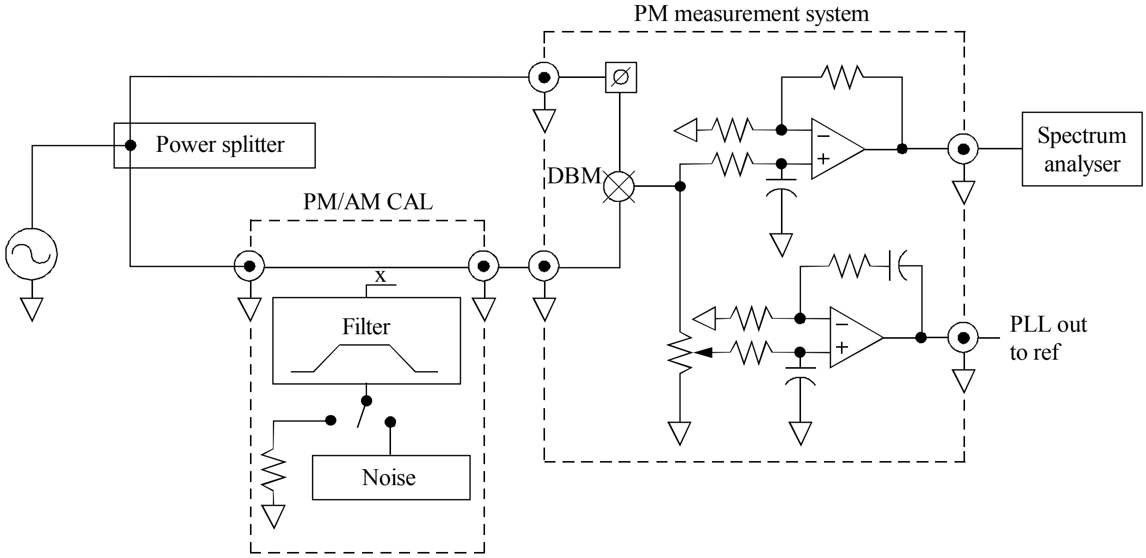


FIGURE 4.12

Block diagram of configuration to measure the noise floor of PM noise measurement systems

Typical double balanced mixers used for phase detection exhibit significant harmonic distortion, especially when driven hard. This feature can be exploited to make PM noise measurements at odd harmonics of the local oscillator or reference frequency. Sensitivities are typically –9 dB for detection at the 3rd harmonic and −16 dB for detection at the 5th harmonic [Walls et al., 1988]. Some specialised mixers are designed to work at even higher harmonics, especially in the microwave and millimetre ranges. This harmonic distortion can cause errors in fundamental PM measurements because the mixer output contains extra terms from the beat between the harmonics of one signal and the noise in the other signal. The errors may be as large as 2-6 dB if one of the signals is a square wave. A low noise lowpass or bandpass filter can be used to eliminate this problem [Walls et al., 1988; Walls et al., 1994].

Another important problem in precision PM measurements is presence of AM noise in sources and measurement systems. Most devices including amplifiers and mixers convert AM noise to apparent PM noise. Typical conversion coefficients are –3 to –30 dB. Therefore to measure the noise floor of a measurement system it is often necessary to use a source with relatively low AM noise [Walls et al., 1988; Nelson et al., 1994; Ascarrunz et al., 1993; Parker, 1989]. The noise floor for many devices is actually set by AM to PM conversion and not the inherent PM noise. References [Walls et al., 1988; Ascarrunz et al., 1993; Parker, 1989] are most revealing in this regard. See Section 4.4 for discussions on techniques for measuring AM noise.

4.2.4 Dual mixer time measurement systems

Figure 4.13 shows a simplified block diagram of a dual mixer time measurement system [Stein et al., 1982; Stein et al., 1983]. This system allows one to measure the phase between a reference oscillator and one or many other sources and is commonly used in system management of clock ensembles. Generally the clocks in an ensemble are run at very nearly the same frequency. Often the fractional frequency offsets are less than 10^{-11}. The standard single heterodyne system can not measure clocks at such low frequency offsets because the data are taken only at the zero crossings and that may happen less than 2 times a day. The dual mixer system circumvents this problem by beating each source against a common offset oscillator that is sometimes phase locked to the reference. By adjusting the offset frequency one can pick the data rate for phase comparison between the various clocks.

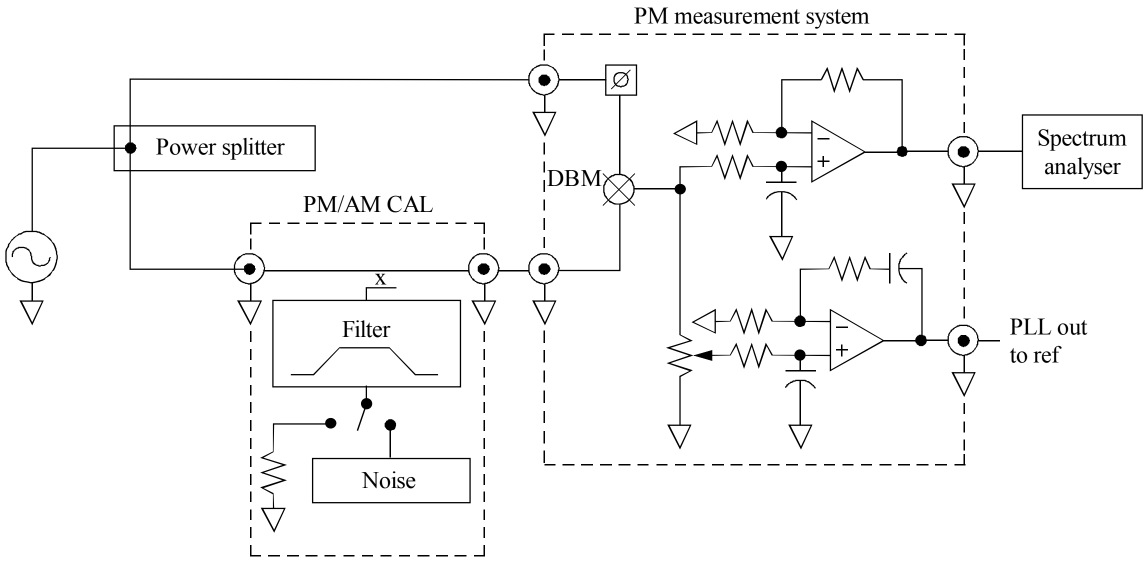


FIGURE 4.13

Block diagram of dual mixer time difference measurement system

This approach is somewhat more complex than single channel systems but has low noise, zero dead time, adjustable resolution, and is able to measure phase, time, frequency, and frequency stability [Stein et al., 1982; Stein et al., 1983]. The fractional frequency resolution of this approach is typically 10^{-12} to 10^{-14} τ^{-1} at a carrier frequency of 5 MHz. This performance should scale as $1/\nu$. The noise floor can be considerably improved using cross-correlation techniques [Lepek et al., 1993; Groslambert et al., 1981]. See Section 4.2.8 for a discussion of the general concepts.

4.2.5 Picket-fence based measurement systems

Figures 4.14 and 4.15 illustrate another method for measuring the frequency stability of precision clocks that also have zero dead-time [Greenhall, 1989]. The reference is divided to some convenient frequency typically of order 10 Hz. The signal starts the time interval counter and a pulse from the divided reference (picket fence) stops the counter. As long as the period of the beat signal is long compared to the reset time of the counter this approach will operate with no dead-time. The details of unfolding the data to recover the phase of the clock are easily accomplished and detailed in [Greenhall, 1989]. This approach is somewhat more complex than single channel systems but has low noise, zero dead time, adjustable resolution, and is able to measure phase, time, frequency, and frequency stability. This approach requires an offset reference to beat against the source under test, similar to the dual mixer time difference system.

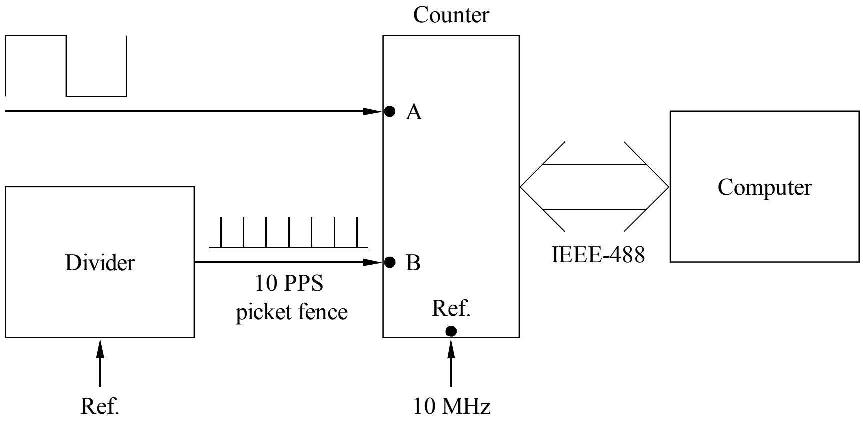


FIGURE 4.14

Block diagram of a picket fence measurement system

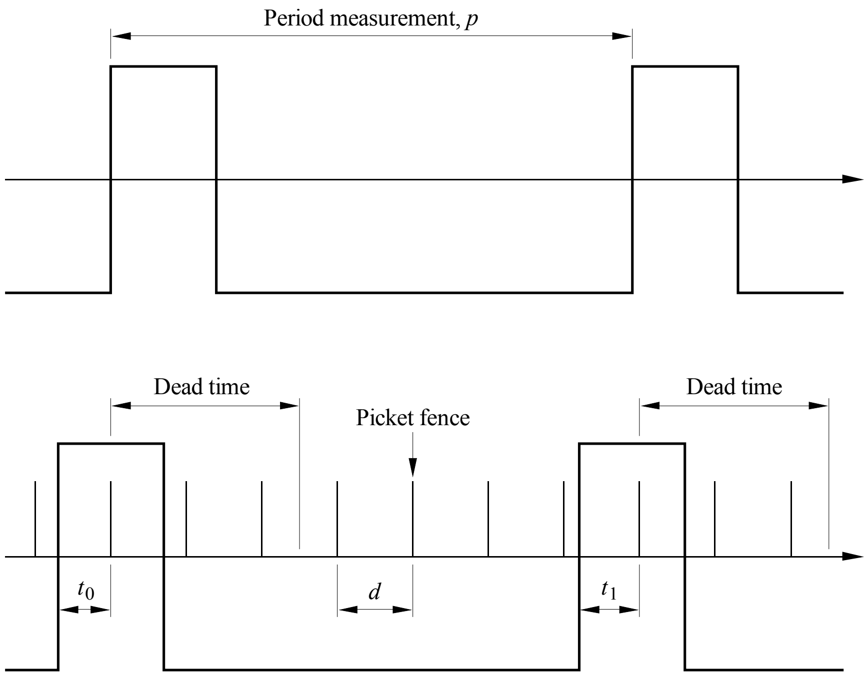


FIGURE 4.15

Timing diagram of a picket fence measurement system

4.2.6 Digital techniques for frequency and PM measurements

If the beat signal between two sources, see Figure 4.5, is digitised with sufficient precision it is possible to recover both time variations of amplitude and phase [Blair, 1994]. The time variations of phase can be analysed to obtain time and frequency as outlined above. The Fourier transform of the zero crossings is proportional to the PM noise. The signal must not saturate the mixer if an output proportional to the amplitude fluctuations of the signal under test is desired. Often the best arrangement for low noise PM measurements is, however, to saturate both the reference and signal ports.

4.2.7 Three-cornered-hat measurements

Simple single channel analogue measurements as described above contain the PM noise of the signal, the reference, and that of the measurement system. Since the noise of the various sources are uncorrelated, the mixer output is proportional to the simple addition of the noise power of all the contributions, that is the various noise components are separable. See Eq.(4.13) [Walls et al., 1988; Walls et al., 1991]. Here β^2 is the AM to PM conversion factor of the mixer.

$$S_{\phi 1,2}(f) = S_{\phi 1}(f) + S_{\phi 2}(f) + S_{\phi meas.syst.}(f) + \beta^2 S_{a1}(f) + \beta^2 S_{a2}(f) \tag{4.13}$$

$$S_{\phi 1}(f) = \frac{1}{2}[S_{\phi 1,2}(f) + S_{\phi 1,3}(f) - S_{\phi 2,3}(f) - S_{\phi meas.syst.}(f) - \beta^2 S_{a1}(f)] \tag{4.14}$$

The noise of the source under test can be sometimes estimated better by measuring the PM noise between three different oscillators. The PM noise of a single oscillator is then estimated from Eq.(4.14). The PM noise floor of the measurement system and the contribution of amplitude noise must be estimated to this method. If the AM noise of the source is higher than the PM noise it may, depending on the value of β, limit the noise floor of the measurement. Typical values of β^2 are from –5 to –30 dB. The noise from the measurement system can be estimated by using the methods discussed in Section 4.2.3. $S_a(f)$ can be measured by the techniques of Section 4.4.

In optimum cases the contribution to $S_{\phi 1}(f)$ due to PM noise from sources 2, 3 and the measurement system can be reduced by a factor of 10. The success of this approach is dependent on carefully matching signal levels at the mixer and measuring $k_d G(f)$ and V_n for each measurement.

4.2.8 Cross-correlation measurement systems

Figure 4.16 shows a two channel measurement system for measuring PM noise of a single source [Walls et al. 1988; Nelson et al., 1994; Ascarrunz et al., 1993; Lance et al., 1984; Walls, 1992]. Each measurement system contains the PM noise of the source under test plus the PM noise of the associated reference signal and the measurement system as described above. Ideally, the average PSD of $(V_{N1} \times V_{N2})$ should contain only the PM noise of the common source. The noise of the two references and the two measurement systems are uncorrelated and should average to zero as the square root of the number of measurements, N. In practice the noise floor is often set by β^2 and the AM noise of the sources.

$$\begin{aligned}\frac{PSD(V_{N1}V_{N2})}{(K_{d1}K_{d2}G_1(f)G_2(f)} &= S_{\phi 1}(f) + \beta^2 (S_{a_1}(f) + S_{a2}(f) + S_{a3}(f)) \\ &+ \frac{S_{\phi}(f)_2 + S_{\phi}(f)_3 + S_{\phi}(f)_{meas.syst.}}{\sqrt{N}}\end{aligned} \tag{4.15}$$

Typically the contribution of PM noise from the sources and the measurement systems can be reduced by a factor of 100 with 10^4 averages. This makes it possible to measure the PM noise in a source that is much better than any other available references. This technique works from the RF region to the mm region [Walls et al., 1988; Nelson et al., 1994; Ascarrunz et al., 1993; Lance et al., 1984; Walls, 1992]. The typical noise floors for this approach for carrier frequencies from 5 MHz to 10.6 GHz are shown in the middle section of

Table 4.7. This cross-correlation approach to 3-cornered-hat measurements has lower noise than that described in Section 4.2.7 because the data is taken simultaneously, which gives better rejection of the noise,

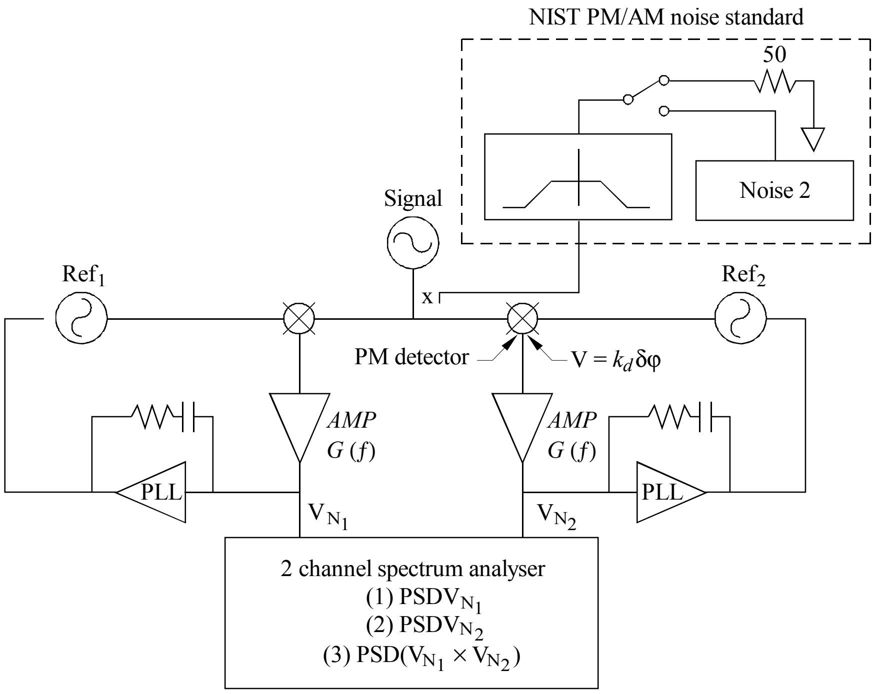


FIGURE 4.16

Block diagram of a two channel measurement system that uses cross-correlation tehniques to estimate the PM noise of a single oscillator

and the results depend linearly on $(k_d\, G(f))^2$ in each channel instead of the difference between them. See Figure 4.17 for a comparison of typical noise floors for different PM measurement techniques [Walls et al., 1988].

Cross-correlation can be used with most single channel techniques to improve the resolution. [Lance et al., 1984] also illustrates the use of cross-correlation to reduce the noise floor of delay line based PM noise measurements. [Lepek et al., 1993; Groslambert et al., 1981] illustrate the use of cross-correlation techniques to reduce the noise floor in time-domain measurement systems.

4.3 Single oscillator measurements of frequency and PM noise

There are a number of techniques that can be used to measure the performance of a single source. All of these techniques are characterised by the use of a delay line or stable resonator as a frequency discriminator. In such systems the output signal is proportional to a frequency difference from nominal, instead of a phase difference between two sources as described in Sections 4.2.0 to 4.2.6. The near carrier noise floor for PM measurements is typically many orders of magnitude worse than the two oscillator techniques outlined in Sections 4.2.0 to 4.2.8 above [Lance et al., 1984; Ashley, 1968]. See Figure 4.17 for a comparison of single oscillator noise floors with that obtained with two oscillator methods. Two channel cross-correlation techniques can be used to improve the noise floor of the single channel measurements shown in Figure 4.17. Another disadvantage of all single oscillator measurement techniques is that it is difficult to verify a noise floor which is better than the testing source.

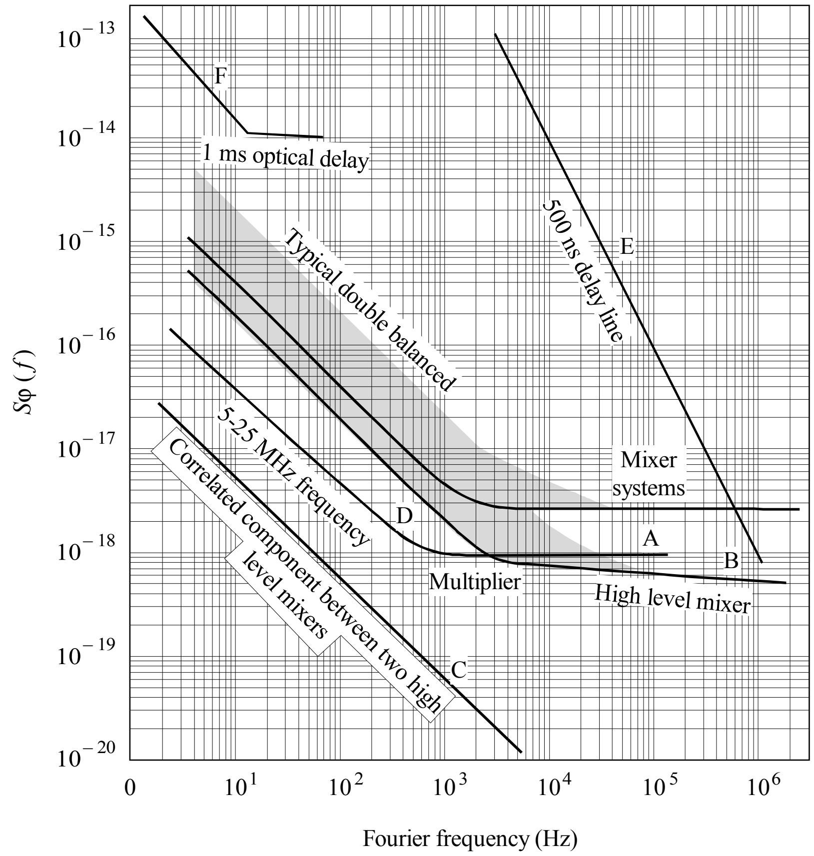


FIGURE 4.17

Comparison of noise floor or resolition of various PM measurement configurations

4.3.1 Delay line measurements of frequency and PM noise

Figure 4.18 shows a block diagram of a simple system to measure frequency and PM noise of a source using a delay line frequency discriminator [Walls et al., 1988]. More elaborate versions are discussed in [Lance et al., 1984; Ashley, 1968; Avramov et al., 1994]. The mixer provides an output voltage proportional to the difference between the phase of the prompt signal and the one from the delay line. The phase shift ϕ is adjusted to yield approximately 0 volts output at a specific frequency. Small changes in phase about this initial operating point are then proportional to small frequency changes. The uncertainty for determining frequency with this approach is quite high, however, it is often very useful for measuring changes in frequency. A simple technique for calibrating the value of $k_v\, G(f)$ is to measure dV, the change in dc output voltage when the frequency of the input signal is stepped by dν which is large compared to the noise and smaller than approximately $1/50\tau$. See Eq. 4.16. The data can be analysed to yield PM noise using Eq. 4.17.

The sensitivity of this approach has a null whenever the phase delay approaches $(2n + 1)90^{o}$ [Ashley, 1968]. Long delay times which are necessary to achieve high sensitivity at low Fourier frequency offsets lead to high attenuation of the signal for RF and microwave signals. Reference [Avramov et al., 1994] reports a noise floor approaching –190 dBc/Hz several MHz from a 1 GHz carrier for a 2-channel delay system with cross-correlation. This problem can be alleviated to some extent by encoding the RF or microwave signal on an optical signal which is matched to a low-loss optical fibre delay line. This improves the sensitivity near the carrier at the expense of the wideband noise floor which is relatively poor due to the noise in optical modulator/demodulator systems. Figure 4.18 of [Lance et al., 1984] compares the data for several different delays against results from the two oscillator method [Avramov et al., 1994].

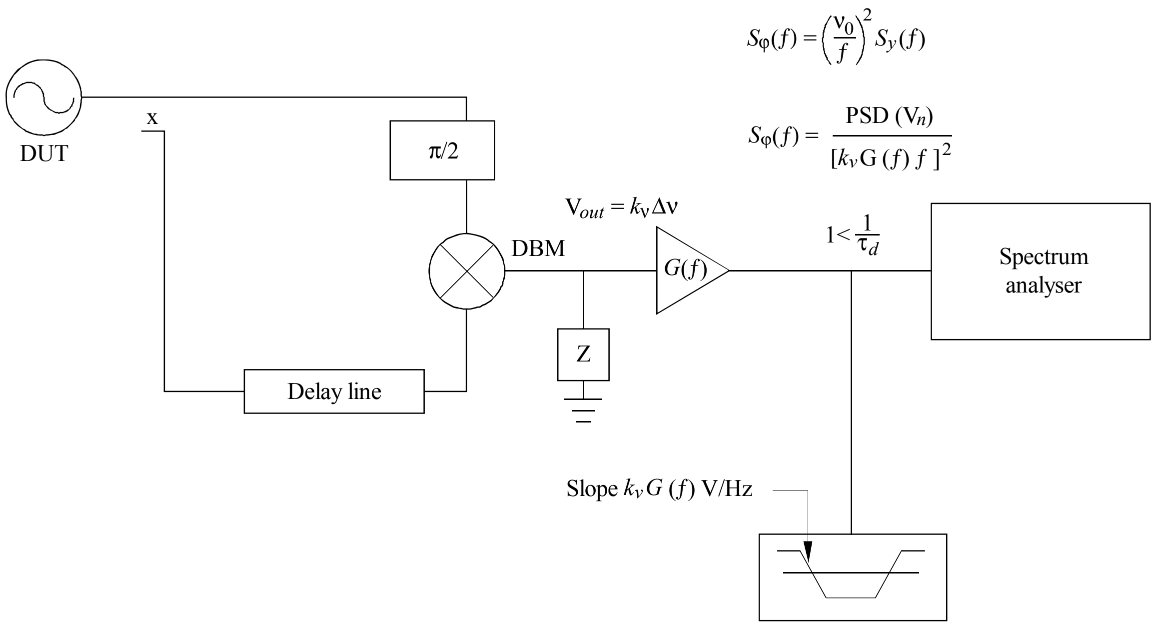


FIGURE 4.18

Block diagram of delay line discriminator for PM measurements

$$dV = k_\nu G(f) \quad d\nu \tag{4.16}$$

$$S_\phi(f) = \frac{V_N^2}{(f k_\nu G(f))^2 BW} \tag{4.17}$$

4.4 Measurements of AM noise

AM noise in sources and devices and AM to PM conversion often limit the performance of practical systems. No PM measurement is complete without a measurement of the AM noise level. The measurement of amplitude modulation (AM) noise can be made without the need for a reference, therefore all measurement techniques are single oscillator measurements.

Figure 4.19 shows a simple single channel measurement system for AM measurements [Walls et al., 1988]. The sensitivity versus Fourier frequency can be calibrated using a source that can be amplitude modulated [Nelson et al., 1992]. The sensitivity of AM systems can also be calibrated using an AM noise standard [Walls, 1993-2]. The advantage of this scheme is that it can usually provide lower uncertainty over a much wider range of Fourier frequencies. Measurements using the AM noise standard in [Walls, 1993-2] provides calibration of K_a $G(f)$ for f from near dc to 10% of the carrier or 1 GHz, whichever is lower.

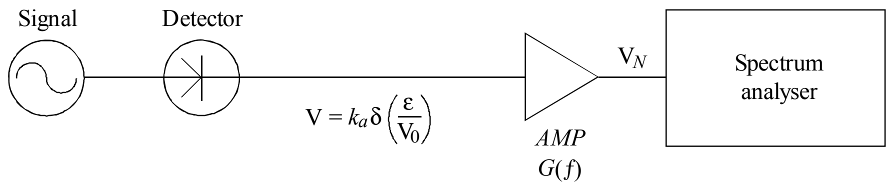


FIGURE 4.19

Block diagram of simple AM noise measurement system

A two channel AM noise measurement system that uses cross-correlation techniques to substantially improve the noise floor can be used to reduce the contribution that the measurement system noise makes to the total measured noise [Walls, 1993-1; Walls, 1993-2; Walls et al., 1994; Nelson et al., 1994; Ascarrunz et al., 1993]. This system is calibrated by the same techniques outlined above.

Tables 4.3 and 4.4 list the parameters which commonly affect the uncertainty of AM noise measurements.

References

ALLAN D.W. and BARNES J.A. [1981] A Modified Allan Variance with Increased Oscillator Characterization Ability. Proceedings 35th Annual Symposium on Frequency Control.

ALLAN D.W., HELLWIG H., KARTASCHOFF P., VANIER J., VIG J., WINKLER G.M. R. and YANNONI N. [1988] Standard Terminology for Fundamental Frequency and Time Metrology. Proceedings 42nd Annual Symposium on Frequency Control.

ASCARRUNZ F.G., FERRE E.S. and WALLS F.L. [1993] Investigations of AM and PM Noise in X-Band Devices. Proceedings 1993 Frequency Control Symposium, Salt Lake City, UT, pp. 303-311.

ASHLEY J.R., SEARLES C.B. and PALKA F.M. [1968] IEEE Transactions Microwave Theory Tech. MTT-16, pp. 753-760.

AVRAMOV I.D., WALLS F.L. and PARKER T.E. [1994] Extremely low thermal Noise Floor Power Oscillator Using Surface Transverse Waves. Proceedings IEEE Frequency Control Symposium.

BARNES J.A. and ALLAN D.W. [1990] Variances Based on Data with Dead Time Between the Measurements. NIST Technical Note 1318.

BARNES J.A., CHI A.R., CUTLER L.S., HEALEY D. J., LEESON D.B., McGUNIGAL T.E., MULLEN J.A. Jr., SMITH W.L., SYDNOR R.L., VESSOT R.F.C. and WINKLER G.M.R. [1971] Characterization of Frequency Stability. *IEEE Transactions on Instrumentation and Measurement*.

BERNIER L.G. [1987] Theoretical Analysis of the Modified Allan Variance. Proceedings 41st Annual Frequency Control Symposium, pp. 116-121.

BLAIR D. [1994] Cryogenic Sapphire Dielectric Resonator Microwave Oscillator with Improved Stability. Proceedings 8th European and Time Forum, Weihenstephan, Germany.

GREENHALL C.A. [1989] A Method for Using a Time Interval Counter to Measure Frequency Stability. *IEEE Transactions. on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. UFFC-36, no. 5, pp. 478-480

GROSLAMBERT J., FEST D., OLIVIER M. and GAGNEPAIN J.J. [1981] Characterization of Frequency Fluctuations by Cross-Correlation and by using 3 or more Oscillators. Proceedings 35th Annual Frequency Control Symposium, pp. 113-118.

HOWE D.A., ALLAN D.W. and BARNES J.A. [1981] Properties of Signal Sources and Measurement Methods. Proceedings 35th Annual Symposium on Frequency Control.

HOWE D. [1995] An Extension of the Allan Variance which Removes a Bias at Long Term. Proceedings International IEEE Frequency Control Symposium.

LANCE A.L., SEAL W.D. and LABAAR F. [1984] Infrared and Millimetre Waves, Vol 11, 239-289.

LEPEK A. and WALLS F. L. [1993] Cross Correlation Analysis Improves Time-Domain Measurements. Proceedings IEEE Frequency Control Symposium, Salt Lake City, UT, pp. 312-320.

LESAGE P. and AUDOIN A [1973] Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements. *IEEE Transactions on Instrumentation and Measurement*, IM-22, pp. 157-161.

LESAGE P. and AYI T. [1984] Characterization of Frequency Stability: Analysis of the Modified Allan Variance and Properties of Its Estimate. *IEEE Transactions on Instrumentation and Measurement*.

NELSON L.M. and WALLS F.L. [1992] Environmental Effects in Mixers and Frequency Distribution Systems. Proceedings IEEE Frequency Control Symposium, Hershey, PA., pp. 831-837.

NELSON L.M., NELSON C.W. and WALLS F.L. [1994] Relationship of AM to PM Noise in Selected RF Oscillators. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*.

PARKER T. E. [1989] Residual Noise in Components. Proceedings 43rd Annual Symposium on Frequency Control, pp. 349-359.

PERCIVAL D. B. and WALDEN A.T. [1993] Spectral Analysis for Physical Application, Cambridge University Press.

RUTMAN J. and WALLS F.L. [1991] Characterization of Frequency Stability in Precision Frequency Sources. *Proceedings IEEE*, Vol. 79, 952-960.

STEIN S., GLAZE D., LEVINE J., GRAY J., HILLIARD D. and HOWE D. [1982] Performance of an Automated High Accuracy Phase Measurement System. Proceedings 36th Annual Symposium on Frequency Control, pp. 314-320.

STEIN S., GLAZE D., LEVINE J., HILLIARD D., HOWE D. and ERB L. [1983] Automated High Accuracy Phase Measurement System. *IEEE Transactions on Instrumentation and Measurement,* IM-32, pp. 227-231.

STEIN S.R. [1992, 1985] Frequency and Time – Their Measurement and Characterization. *Precision Frequency Control*, Volume 2, Nelson et al.

STEIN S.R. and TURNEEAURE J.P. [1973] The Development of the Superconducting Cavity Stabilized Oscillator. Proceedings 27th Annual Symposium on Frequency Control, pp. 414-420.

SULLIVAN D.B., ALLAN D.W., HOWE D.A. and WALLS F. L. [1990] Characterization of Clocks and Oscillators. NIST Technical Note 1337, pp. 1-342.

TAYLOR B.N. and KUYATT C.E. [1993] NIST Technical Note TN 1297.

WALLS F.L. [1990] Method and Apparatus for Wide Band Phase Modulation, United States Patent – No.4, pp. 968, 908.

WALLS F.L. [1992] Frequency Calibration Standard Using a Wide Band Phase Modulator, United States Patent – No. 5, 101, 506.

WALLS F.L. [1993] Reducing Errors. Complexity and Measurement Time for PM Noise Measurements. Proceedings Frequency Control Symposium, Salt Lake City, UT, pp. 81-86.

WALLS F.L. [1993] Secondary Standard for PM and AM Noise at 5, 10 and 100 MHz. *IEEE Transactions on Instrumentation and Measurement*, IM-42, pp. 136-143.

WALLS F.L., BARILLET R., BESSON R., GROSLAMBERT J., SCHUMACHER P., RUFENACHT J. and HILTY K. [1994] International Comparison of Phase Noise. Proceedings 8th European Frequency and Time Forum, Weihenstephan, Germany.

WALLS F.L., CLEMENTS A.J.D., FELTON C.M., LOMBARDI M.A. and VANEK M.D. [1988] Extending the Range and Accuracy of Phase Noise Measurements. Proceedings 42nd Annual Symposium on Frequency Control, pp. 432-441.

WALLS F.L., CLEMENTS A.J.D., FELTON C.M. and MARTIN T.D. [1991] Precision Phase Noise Metrology. Proceedings National Conference of Standards Laboratories (NCSL), Albuquerque, NM, August, pp. 257-275.

WALLS F. L. and DEMARCHI A. [1975] RF Spectrum of a Signal After Frequency Multiplication, Measurement and Comparison with a Simple Calculation. *IEEE Transactions on Instrumentation and Measurement,* IM-24, pp. 210-217.

WALLS F. L., GARY J., O'GALLAGHER A., SWEEET L. and SWEET R. [1990] Time-Domain Frequency Stability Calculated from the Frequency Domain: an Update. Proceedings 4th European Frequency and Time Forum, Neuchatel, Switzerland, pp. 197-204.

WALLS F.L., PERCIVAL. D.B. and IRELAN W.R. [1989] Biases and Variances of Several FFT Spectral Estimators as a Function of Noise Type and Number of Samples. Proceedings 43rd Annual Symposium on Frequency Control, pp. 336-341.

WALLS F.L. and STEIN S.R. [1976] Servo Techniques in Oscillators and Measurement Systems. NBS Technical Note 692, pp. 1-20.

WALLS W.F. [1992] Cross-correlation Phase Noise Measurements. Proceedings 1992 Frequency Control Symposium, pp. 257-261.

WALTER T. [1994] Characterizing Frequency Stability: A Continuous Power-Law Model with Discrete Sampling. *IEEE Transactions on Instrumentation and Measurement*, Vol. 43, 69-79.

WEISS M.A., WALLS F.L., GREENHALL C.A., WALTER T. [1995] Confidence on the Modified Allan Variance. Proceedings 9th European Frequency and Time Forum, Besançon, France.

Chapter 5

Characteristics of various frequency standards

Contents

5.1 Definitions and discussion: measures and implications

5.1.1 The characterisation of random processes

5.1.1.1 $\mathcal{L}(f)$, $S_\phi(f)$

$S_\phi(f)$ is the power spectral density of phase (see Chapters 2 (Part A), 3 and 4). It is what one would see on a spectrum analyser that only measured phase noise. It is important in applications where the short-term stability is critical such as high frequency communications and where noise of timing signals must be very small. $\mathcal{L}(f)$ is defined as 1/2 $S_\phi(f)$ or, in dB, 3 dB lower than $S_\phi(f)$. The meaning of the two definitions is that $\mathcal{L}(f)$ is the one-sided density and $S_\phi(f)$ is the two-sided density (this is not strictly true for oscillator frequencies very close to the carrier.) The units of both are dBc/Hz; that is, the amount of phase-noise power in a bandwidth of one Hertz, relative to the power in the carrier. Phase noise on a signal increases with frequency multiplication by 20 log N, where N is the frequency multiplication factor. Therefore, for use of a frequency source at the higher frequencies, say 30 GHz, $S_\phi(f)$ must be very small at the base frequency in order to maintain a low noise capability in either transmitting a signal or receiving one. For example, if one requires a phase noise level such that $\mathcal{L}(f)$ is less than –50 dBc at 30 GHz, then the basic signal, say at 5 MHz, must have a $\mathcal{L}(f)$ of $-50\ -20\ \log(30 \times 10^9\ /5 \times 10^6)$, or $-40\ -20\ \log(6 \times 10^3) = -50 - 75.6 = -125.6$ dBc/Hz. This value requires a high quality precision quartz oscillator and is not usually obtainable from other precision frequency sources. Similarly, when making precise time measurements, a low value of $\mathcal{L}(f)$ allows the measurement of time intervals with better accuracy. Whenever $\mathcal{L}(f)$ or S are given, the carrier frequency must be given; otherwise, the specification is not completely defined.

5.1.1.2 $\sigma_y(\tau)$, mod $\sigma_y(\tau)$, $\sigma_x(\tau)$

These measures are similar to, and derived from, the numbers that a frequency counter would measure. The Allan deviation, $\sigma_y(\tau)$, is derived from the difference in the values obtained from a series of adjacent measurements of a frequency source (see Chapters 2 and 4). Since $\sigma_y(\tau)$ is not capable of detecting the difference between flicker of phase and white phase noise processes, mod $\sigma_y(\tau)$ was developed in order to more completely describe the noise processes in the frequency source. These measures are useful for describing the performance of the frequency source when used as a timing generator or clock. The measure $\sigma_x(\tau)$ was developed in characterising time stability. It is defined in Chapter 4. This measure is very useful in characterising the random variations in measurement systems, time distribution systems, and networks. The communication industry has adopted it as a measure. Its construction is related to that of mod $\sigma_y(\tau)$, note that $\sigma_x(\tau) = \tau(\mathrm{mod}\,\sigma_y(\tau)/\sqrt{3})$ and is usually derived from the phase or time differences rather than frequency differences as described in Chapter 3.

5.1.2 Systematic effects

5.1.2.1 Environmental effects

5.1.2.1.1 Temperature [IEEE, 1995]

All frequency sources are sensitive to temperature to a greater or lesser extent. Depending on this sensitivity and the performance required, special precautions may be required in order to control the temperature of the local environment. For example, if the temperature sensitivity of a frequency source is 1×10^{-12}/°C and the room in which the frequency source is to be installed has a temperature variation of 2° C peak-to-peak from the air-conditioning equipment with a period of 5 minutes, the frequency variation will be approximately 2×10^{-12} with the same period. (Note that the actual amplitude of the variation is affected by the thermal time constant of the frequency source and the variation may be reduced by increasing the time constant.) If the requirement is that the frequency source be stable to 1×10^{-13} at 300 seconds, the stability of the room (or separate enclosure) must be improved by a factor of 20, to 0.1° C peak-to-peak. An excellent room air conditioning system can achieve a temperature stability of 2° C peak-to-peak, so some extra temperature control may be necessary in this example. Note that precision air conditioning systems can be custom designed to have a temperature variation no larger than ±0.05° C for a normal room.

5.1.2.1.2 Humidity [IEEE, 1995]

Changes in the ambient humidity can affect the frequency source in several ways. The thermal capacity and conductance of moist air is different from that of dry air. This can affect the internal thermal environment of the frequency source and the thermal gradients as well as changing the resistance of the printed circuit boards in high impedance circuits. All these effects may change the frequency of the frequency source. Some earlier cæsium frequency sources exhibited very strong influence by humidity, but this is greatly reduced in the newer sources. The result of humidity dependence is typically seen as a seasonal change in the frequency of the source, or a change with a period of approximately 4½ days, the average period due to weather fronts moving through an area, world-wide.

5.1.2.1.3 Barometric pressure [IEEE, 1995]

Since the density of the air changes with changing barometric pressure, the thermal characteristics of the frequency source are affected, similarly to the effects of humidity. In addition, flexure of various parts of the frequency source may cause frequency changes, e.g. the cavity in a Hydrogen Maser. These changes are usually related to the moving of weather fronts through an area with an average period, again, of 4½ days, world-wide.

5.1.2.1.4 Magnetic field [IEEE, 1995]

All frequency sources are sensitive to magnetic fields to some degree. Similar considerations must be applied as in the case of temperature variations. Magnetic shielding is often required. The exact placement of a frequency source in an electronics rack relative to other permeable items can have a considerable effect. Magnetic fields of as much as 5 µT (50 mgauss) at the power line frequency (50 or 60 Hz) are usually found in a typical electronics rack. Static or slowly varying fields due to the movement of automobiles or other large steel objects can affect the frequency source, as can variations of the earth's field due to temperature changes of the reinforcing steel in the building which affects the permeability of the steel, and thus the magnetic field in the environment of the frequency source. Location of the frequency source is very important in order to reduce these effects to an acceptable level. Degradation of the performance due to power line related fields can be readily ascertained by the use of a good spectrum analyser as described in Chapter 3. Slow changes can be discovered only by a continuous monitoring of the frequency of the source with respect to an external reference, such as GPS or a national laboratory as described in Chapters 2 (Part B) and 6. The internal magnetic shields on frequency standards are also non-linear, that is, they have different shielding coefficients for different levels of magnetic fields. They also have a certain amount of hysteresis, exhibiting a memory of past events in terms of their shielding coefficients. The frequency standard itself has different susceptibility to magnetic fields imposed in different directions, so that location and orientation are very important.

5.1.2.1.5 Power line voltage, noise, and interruptions [IEEE, 1995]

Power supply line voltage variations may affect the frequency of the frequency source. Noise on the power line may modulate the frequency source or be added to the output. Interruptions of the power line voltage due to storms or failure will not only stop the operation of the source, but may result in a long time while the source is not within specification while operation re-stabilises (a month or longer for some frequency sources). It is good practice to operate the frequency source from an uninterruptible power system (UPS). In many cases, batteries are adequate, at least for short outages. Generators may be necessary to maintain operation of the frequency source for the longest expected outage. Stable, low noise power supplies (preferably redundant) are necessary to reduce or eliminate the effects of power line noise, surges, and drop-outs and droops.

5.1.2.1.6 Acceleration, vibration, and shock [IEEE, 1995]

Any acceleration of the frequency source causes stresses on the internal parts of the source. The resulting flexure may cause the frequency to change. Some of the parts may also exhibit magnetostrictive properties which can affect the operation of the source and change the frequency. Quartz crystal oscillators are especially sensitive to these effects. Since all frequency sources include a quartz crystal oscillator, care must be taken to prevent vibration from being large enough to modulate the output frequency of the source. A typical office or laboratory has an vibration of ≈ 0.2 m/s^2 due to air conditioning, etc. Since the vibration sensitivity of a quartz oscillator lies in the range of 10^{-10} to 10^{-11} per g, an evaluation of the necessary level of vibration reduction must be performed. In atomic frequency sources, the quartz oscillator is responsible for the short-term performance. The atomic resonator or oscillator controls the performance in the medium and long term. The actual cross-over between the quartz oscillator and the atomic resonator depends on the type of standard. For the hydrogen maser it is at approximately 0.5 Hz, while for the cæsium beam standard, it is on the order of 0.01 Hz.

Shocks can be large enough to permanently affect the frequency of the source by displacing the elements in the quartz resonator or in the atomic device. These frequency sources are to be treated as fragile devices during shipping and handling.

5.1.2.1.7 Ageing [IEEE, 1995]

Many frequency sources produce a frequency that, under fixed environmental conditions, changes with time, usually in a quasi-linear fashion. This affect is called ageing. It is caused by relaxation of stresses in the resonant element or in an element that is closely coupled to the resonant element. Frequency sources that typically exhibit this behaviour are: Quartz crystal oscillators, due to changes in the quartz crystal itself and also in the associated circuitry; Hydrogen Masers due to changes in the cavity and, possibly, changes in the wall coating. Cæsium usually does not exhibit ageing at a discernible level until they are near the end of life of the beam tube. Rubidium standards can exhibit ageing, for example, due to the reaction of the rubidium gas with the glass in the lamp and the gas cell.

5.1.2.1.8 Drift [IEEE 1995]

The meanings of the terms “ageing” and “drift” are often confused. The accepted definition of drift is the long-term change in frequency due to all causes, including systematic terms and random terms, including ageing. Ageing, as explained above, is the frequency change due to internal effects with the standard operating in a fixed environment.

5.2 Characteristics of various frequency sources [CCIR, 1990]

Figures 5.1 and 5.2, and Table 5.1 demonstrate the behaviour of various frequency sources in a constant environment. In particular, note that the cæsium standard has the best long term stability as behooves its original choice for the definition of the second. While the standards depicted in these Figures are all commercial standards, most operational day-to-day national and calibration laboratories use them, availing themselves of true primary standards (in the case of national laboratories) only for calibration purposes or for calibrating the length of the UTC second for the BIPM.

Figure 5.1 shows the power spectral density of phase for the quartz frequency standard, the rubidium frequency standard, the cæsium frequency standard and the hydrogen maser frequency standard. Note that, for the higher Fourier frequencies (farther from the carrier), they all approach the value for the quartz frequency standard. This is because they all incorporate a quartz oscillator as part of the standard and as an output device, see Chapters 1 and 2 (Part A). The best commercial standard for the lower Fourier frequencies is the hydrogen maser, followed in order by the cæsium beam frequency standard, the rubidium standard and the quartz oscillator.

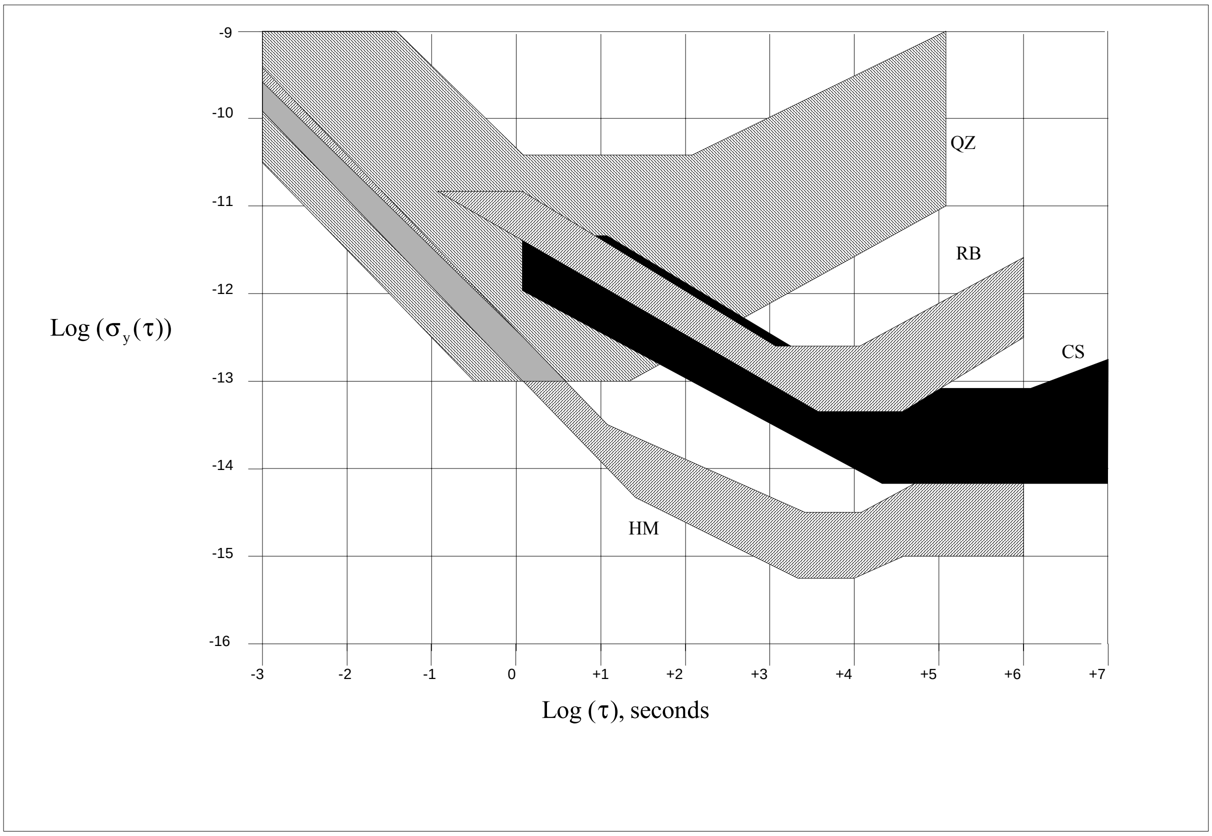


FIGURE 5.1

Stability Ranges of Various Frequency Sources for 1 kHz

Figure 5.2 shows the Allan deviation for the same frequency standards. Again, for the shorter averaging times, they all approach the performance of the quartz oscillator which is an integral part of their systems. In the very long term, the cæsium is the best (the hydrogen maser with auto-tuning approaches the cæsium performance, but it is not a primary standard). The quartz, rubidium and hydrogen standards all exhibit a random walk of frequency in the long term. The best cæsiums do not exhibit this behaviour until nearly the end of life of the cæsium beam tube.

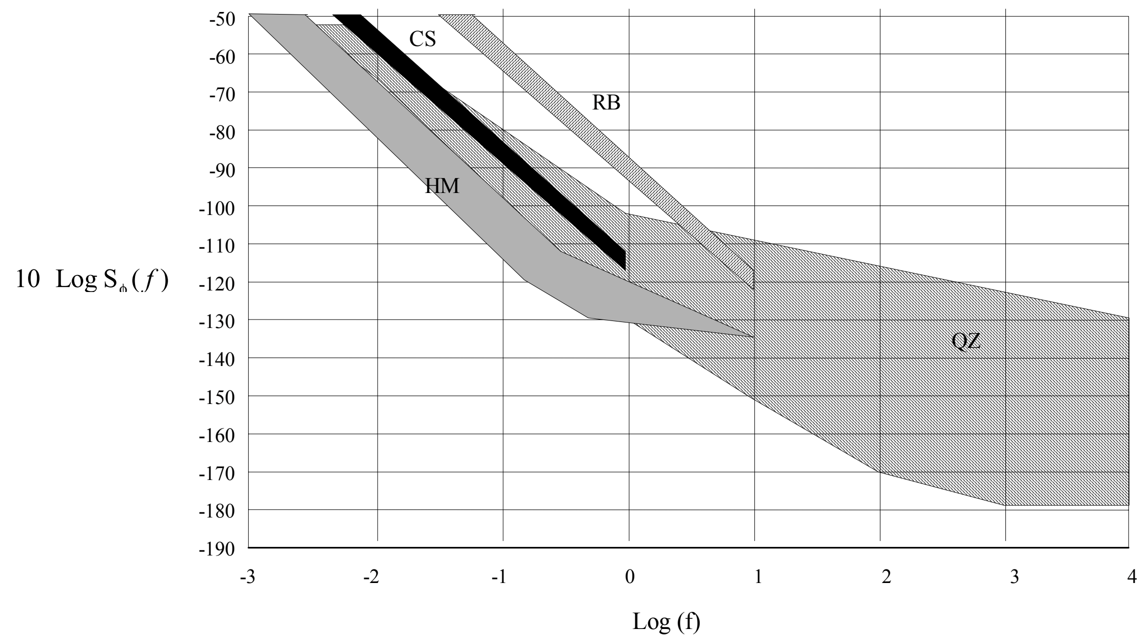


FIGURE 5.2

Nominal Power Spectral Density of Phase for Various Standards at 5 MHz

Table 5.1 lists the important physical parameters of the various standards. These include the frequency uncertainty, a condensation of the performance , the volume, weight power requirements and estimated cost. A quick survey of this list demonstrates the trade-off between cost, volume, weight and performance.

Table 5.2 lists the typical environmental sensitivities of the quartz, hydrogen maser, caesium and rubidium standards.

Figure 5.1 shows the range of Allan deviation performance available for the various frequency sources [CCIR, 1990] . The difference in performance from the top of one of the bands to the bottom of the band depends on several factors including moderness of the design, cost of the standard, and stability of the environment.

Figure 5.2 shows the range of the power spectral density of phase, $S_\phi(f)$, for the various standards [CCIR, 1990]. The same description of the meanings of the widths of the bands applies as in Figure 5.1.

TABLE 5-1

Frequency Source Performance in controlled Environment

Frequency Standard	Uncertainty	Stability			Volume (dm^3)	Mass (kg)	Power Requirement	Estimated Cost, 1987 (US k$)
		Short Term (100 s)	Floor	Ageing, per year				
Precision quartz	(1)	10^{-10} to 10^{-13} (2)	10^{-10} to 10^{-13}	10^{-6} to 10^{-10}	1 to 10	0.1 to 10	0.1 to 20	0.5 to 20
Hydrogen Maser	10^{-12}	$2 \rightarrow 6 \times 10^{-15}$	8 to 20×10^{-16}	10^{-12} to 10^{-14}	1000	250	100	200 to 450
Commercial Cæsium								
(3)	2×10^{-12}	10^{-12}	$2 \rightarrow 5 \times 10^{-14}$	$<3 \times 10^{-13}$	45	30	30	40 to 75
(4)	1×10^{-12}	8.5×10^{-13}	$3 \rightarrow 5 \times 10^{-15}$	$<3 \times 10^{-13}$	45	30	30	67 to 75
Rubidium, (high performance)	(1)	7×10^{-15}	4×10^{-14}	10^{-10}	26	15	35	20

(1) The specification does not apply.
(2) Stability at 1 second.
(3) "High performance Unit," early units [SYDNOR, 1989]
(4) "High performance Unit," late units [KUSTERS, 1992]

TABLE 5-2

Environmental Sensitivities of Various Frequency Sources

Frequency Standard Type	Temperature, per K	Acceleration, per m/s^2*	Magnetic Field, per Tesla	Barometric Pressure, per Pascal	Ageing, per year
Precision Quartz, Oven Controlled	10^{-12}	10^{-11}	10^{-10}	10^{-12}	10^{-8}
Hydrogen Maser	10^{-14}	10^{-14}	10^{-10}	10^{-12}	10^{-12}
Cæsium Beam					
(3)	10^{-14}	10^{-14}	10^{-9}	10^{-13}	10^{-12}
(4)	$\leq 1 \times 10^{-15}$	10^{-14}	$\leq 1 \times 10^{-14}$	$\leq 1 \times 10^{-15}$	$< 2 \times 10^{-14}$
Rubidium Cell	10^{-12}	10^{-13}	10^{-13}	10^{-15}	10^{-10}

* For frequencies inside the servo bandwidth. Outside the bandwidth, this sensitivity is that of the quartz oscillator.
(3) "High performance Unit," early units [SYDNOR, 1989]
(4) "High performance Unit," late units [KUSTERS, 1992]

References

CCIR [1990] Reports. Annex to Volume VII, Report 364-6. Performance of standard-frequency generators, pp. 129-149

IEEE STANDARD 1193 [1995] IEEE Guide for Measurement of Environmental Sensitivities of Standard Frequency Generators.

KUSTERS J.A. and JOHNSON J.L. [1992] A New Cæsium Beam Frequency Standard Performance Data. Procedings 1992 IEEE Frequency Control Symposium, pp. 143-150.

SYDNOR R.L., TUCKER T.K., GREENHALL G.A., DIENER W.A. and MALEKI L. [1989] Proceedings Twenty-first Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, pp. 409-419.

Chapter 6

Time Scales

Contents

6.1 Introduction

The true nature of time has no rational explanation; we simply feel that time never stops or reverses. But, apart from philosophical considerations, experience indicates that any event can be localised by specifying three space coordinates and one temporal coordinate, generally denoted (x, y, z, t). So far, no experiment has ever called for more independent parameters. Intuitively, a time scale is thus defined as the time axis of a system of coordinates. Recommendation ITU-R TF.686 defines a time scale as an *ordered set of scale markers with an associated numbering*.

At first sight, establishing a time scale seems a simple task, as any evolving system allows the transformation of the measure of time variation into the measure of a dimensional quantity. But scientists ask for metrological properties, and reliable, stable, and accurate reference time scales are required. In addition, there is no absolute time as conceived by Newton in classical mechanics. In reality, all temporal phenomena are affected by gravitational fields and velocities with respect to the observer. Time scales must thus be defined in the framework of general relativity, as explained in Section 6.2.

Conventionally, one distinguishes two different types of time scale: integrated time scales and dynamic time scales:

- For integrated time scales, the primary data is a unit of duration, that is, of time interval, defined from a physical phenomenon. The time scale is constructed by fixing a conventional origin and by accumulating units of duration without dead-time and without interruption. This concept for the construction of a time axis was applied to the duration of the day, leading to the definition of Mean Solar Time. The present worldwide reference time scale, International Atomic Time, TAI, is an integrated time scale; it is obtained by the accumulation of atomic seconds defined as a number of periods of the radiation corresponding to a given transition of the caesium atom.

- For dynamic time scales, the primary data results from the observation of a dynamic physical system, described by a mathematical model in which time is a parameter that unambiguously identifies the configuration of the system. The time measurement thus becomes a position measurement, and the unit of time is defined as a particular duration. Universal Time, UT1, and Ephemeris Time, ET, are dynamic time scales, based respectively on the rotation of the Earth on its axis and around the Sun.

In the past, a number of time scales, either dynamic or integrated, have been defined. The associated unit of duration was then used to define the second of the International System of Units (SI). The change from one definition to another has been motivated by the desire to improve accuracy. A brief summary is given in the following sections.

6.1.1 Universal time

Universal Time, UT1, is a dynamic time, derived from the observation of the Earth’s rotation: it is proportional to the angle of rotation of the Earth on its axis. The coefficient of proportionality is chosen so that 24 hours of UT1 are close to the mean duration of the day, and the phase is chosen so that 0 h UT1 corresponds, on average, to midnight in Greenwich.

The associated unit of time is the second of mean solar time. Its definition is not very precise: it is *the fraction 1/86 400 of the mean solar day*. This was the SI second till 1960. Astronomers estimated that it could be realised with an uncertainty of order 10^{-9}, this level of accuracy being achieved after a decade of astronomical observations, analysis and averaging.

The UT1 was the reference time scale, and thus the basis of legal time, until 1972. It still provides a record of the Earth’s rotation for geophysics, astronomy, *etc.*

6.1.2 Ephemeris Time

Ephemeris Time (ET) is a dynamic time, derived from the theory of the Earth’s rotation around the Sun (including the rotation of the moon around the Earth, *etc.*): it is provided through an expression for the mean longitude of the Sun. This expression and the initial phase of ET were chosen so that ET and UT1 were in approximate coincidence in 1900. Since then they have slowly diverged (ET – UT1 ≈ 56 s in 1988).

The associated unit of time is the ephemeris second, defined as *the fraction 1/31556925,9747 of the tropical year for 1900 January 0 at 12 hours ephemeris time*. This was the SI second from 1960 to 1967.

Through observations of planets and of the moon, it has been possible to obtain the time differences between ET and UT1 since 1630, with a precision of several tens of seconds for the 17th century, and several seconds for the 19th century; ET thus constitutes a reference for studying the Earth's rotation in the past.

6.1.3 International Atomic Time

International Atomic Time (TAI) is an integrated time available since 1955. The unit of time is the atomic second, which became the SI second in 1967 and is still in use. Its definition, adopted by the 13th Conférence Générale des Poids et Mesures in 1967, is as follows:

> *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 caesium 133 atom.*

The atomic second can be realised in a laboratory. The best commercial caesium clocks, now widely used in timing centres, produce it with a stated accuracy of order 1×10^{-12}. Some laboratories maintain a number of primary frequency standards; these give an ultimate accuracy of realisation of order 1×10^{-14} (see Chapter 1).

The function of a clock is to provide a continuous, ordered set of markers with an associated numbering. This constitutes a time scale. The number associated with a given marker is designated as the "reading of the clock". Since physical devices can fail, laboratories are inevitably led to keep several clocks. Clock data are collected and treated together through a combination of their readings in order to generate an ensemble time. On a worldwide scale, such an ensemble time is International Atomic Time, TAI.

The definition of TAI was approved by the Comité International des Poids et Mesures in 1970, and recognised by the Conférence Générale des Poids et Mesures in 1971. It reads as follows:

> *International Atomic Time (TAI) is the time reference coordinate established by the Bureau International de l'Heure on the basis of the readings of atomic clocks operating in various establishments in accordance with the definition of the second, the unit of time of the International System of Units.* [In 1988, the responsibility for TAI was transferred to the Time Section of the Bureau International des Poids et Mesures, BIPM].

The origin of TAI has been agreed officially to coincide with UT1 on 1st January 1958.

An important consequence of improved accuracies in the realisation of the atomic second is that relativistic effects are significant. In this context, the definition of the second must be understood as the definition of proper time, *i.e.*, strictly speaking, the user must be in the neighbourhood of the clock and at rest with respect to it. When comparing two realisations of the SI second, differences of a few parts in 10^{13} may appear due to the different gravitational fields to which the clocks are subjected. The definition of TAI was thus completed as follows, in a declaration of the Comité Consultatif pour la Définition de la Seconde [CCDS Report, 1980] during its 9th Session held in 1980:

> TAI *is a coordinate time scale defined in a geocentric reference frame* (origin of the frame at the centre of the Earth) *with the SI second as realised on the rotating geoid as the scale unit.*

Hence a new situation (unlike UT1 and ET) arose in which the relation between the TAI scale unit and a given realisation of the SI second depends on the position of the clock which produces it. For all clocks fixed on the Earth and situated at sea level, the scale unit of TAI is equal to the unit of time as realised locally; but the scale unit of TAI appears to be longer by 1.1×10^{-13} s when compared with a clock at a 1 000 m altitude, due to the gravitational red frequency shift [Misner *et al.,* 1973]. A complete theoretical definition of TAI, in the framework of general relativity, was given for the first time in 1991 by the International Astronomical Union, IAU, (see Section 6.2).

6.1.4 Coordinated Universal Time

Coordinated Universal Time, UTC, was defined in 1972. It represents a combination of the two time scales TAI and UT1, and is defined by the following system of equations valid for any date t:

$$UTC(t) - TAI(t) = n \text{ seconds } (n \text{ integer}) \qquad (6.1)$$

and

$$|UTC(t) - UT1(t)| < 0.9 \text{ s.}$$

The quantities $UTC(t)$ and $TAI(t)$ differ, for any date t, by an integer number of seconds, equal to 29 from 1st July 1994. The International Earth Rotation Service, IERS, which is responsible for the publication of UT1, decides on the adjustment of seconds by reference to the predicted divergence between the time scales UT1 and TAI. Leap seconds are introduced at the end of a month, normally at the end of June or December.

By definition, UTC has the same metrological properties as TAI, which is an atomic time. In addition, it follows the rotation of the Earth to within 1 second. This is sufficient for the purposes of astronomical navigation, where UT1 is required in real time.

The UTC is the general basis for the distribution of time around the world. Local times are derived from UTC by a shift of an integer number of half-hours (which can change from winter time to summer time), decided by the administration of individual countries, or regional groups. All time signals, at whatever level, including signals distributed by TV, radio, or speaking clocks, are synchronised on these local times, and thus to UTC.

The reference time scales TAI and UTC are calculated and distributed by the Time Section of the BIPM. These are deferred-time time scales, for which to achieve ultimate metrological quality requires months of data collection and treatment. National time laboratories thus keep other time scales, for more immediate use, which are carefully compared *a posteriori* with TAI and UTC at each new issue. These are the independent local time scales, TA(k), and the local representations of UTC, UTC(k), where k designates the acronym of the laboratory.

In 1994, 17 independent TA(k) were maintained. They were generated from ensembles of clocks carefully kept on one site, such as is the American A.1(MEAN) from the USNO and AT1 from the NIST, or on several different sites within the country, such as is done for the French TA(F) and the Swiss TA(CH). The basic measurement cycle is much shorter than for TAI (1 hour to 1 day against 10 days), the time scale is updated much more often (every day or week against every 2 months), and the update may be calculated *a posteriori* or in almost real time. These time scales are free-running and have no physical representations. They are known through values of time differences with respect to a physical clock which is also kept in the laboratory. Their scientific purpose is to provide a stable local reference.

In 1994, 45 UTC(k) were in operation throughout the world. They are generally linked to the output of a clock, with or without frequency correction, and thus correspond to a physical signal accessible in real time. They are not free-running but are closely steered to follow UTC. According to Recommendation ITU-R TF.536 the recommended maximum time difference between the time scales UTC and UTC(k) is ± 1 µs, and the goal is to reach $\pm$ 100 ns [CCDS Report, 1993]. The UTC(k) provide real-time synchronisation, in particular they are used as references for broadcast time signals.

Other time scales support applications in navigation and timing, through satellite global navigation systems. The two principal ones are GPS time, for the American Global Positioning System, and GLONASS time for the Russian GLObal NAvigation Satellite System. Both are generated with a high update rate (of the order of several minutes) from an ensemble of clocks kept in the master control station of the system, and steered on a local representation of UTC: UTC(USNO) for GPS time and UTC(SU) for GLONASS time.

For most demanding applications, such as millisecond pulsar timing, the BIPM issues atomic time scales retrospectively. These are designated TT(BIPMxx) where 1900 + xx is the year of computation [Guinot, 1988]. The successive versions of TT(BIPMxx) are both updates and revisions: they may differ for common dates.

Before dealing with the practice of computing and disseminating time scales such as TAI, UTC, TA(k), UTC(k), GPS time, GLONASS time, and TT(BIPMxx), in Sections 6.3 and 6.4, we return to the theoretical definitions of time scales in general relativity.

6.2 Time scales in general relativity

6.2.1 Coordinate systems in general relativity

In general relativity, time scales are considered as one of the coordinates of four-dimensional space-time reference systems.

Due to the curvature of space-time, the scale units of these coordinates do not, in general, have a globally constant relation to locally measurable (proper) quantities [Misner *et al.,* 1973, Brumberg 1991]. In the framework of Newtonian mechanics, it is always possible to define coordinates in such a way that their scale units are everywhere equal to measured distances and durations. This is impossible in general relativity, where the relation between measured quantities and coordinate scale units depends on the position in space-time of the measuring observer. For time scales, this implies that the relation between a coordinate time interval and the locally realised second, using an atomic clock for example, depends on the position of the clock.

In principle, one is free to use any set of coordinates for the description of space-time. However, it turns out that by defining several overlapping systems of coordinates, each valid in a restricted region, the treatment of practical problems and the relationship between coordinates and measurable quantities can be greatly simplified [IAU, 1992]. Such definitions provide several time coordinates, each valid in a particular region of space-time, with the relation between them given by relativistic coordinate transformations.

A system of coordinates in general relativity is defined by its metric tensor $g_{\alpha\beta}(x\lambda)$ (Greek indices run from 0 to 3) which is position- and time- dependent and must be known for the whole region of space-time within which the coordinate system is used.

The need to define several relativistic systems of space-time coordinates, in particular barycentric and geocentric ones, was recognised by the International Astronomical Union IAU in its 1991 Resolution A4 [IAU, 1991, 1992]. This Resolution includes definitions of barycentric and geocentric coordinate time scales, and so provides the theoretical basis for the definition of TAI.

6.2.2 The 1991 IAU Resolution A4

The International Astronomical Union approved Resolution A4 at its General Assembly held in Buenos Aires in August 1991. The complete text of this Resolution is in the 1992 IAU Information Bulletin 67. It contains several Recommendations of importance for the definition and realisation of coordinate time scales. They are explained in the following sections.

6.2.2.1 Recommendation I

Recommendation I explicitly introduces the general theory of relativity as the theoretical background for the definition of space-time reference frames. It provides the form of the metric to be used for coordinate systems centred at the barycentre of an ensemble of masses:

$$\mathrm{d}s^2 = -c^2\mathrm{d}\tau^2 = g_{\alpha\beta}(x^\lambda)\mathrm{d}x^\alpha \mathrm{d}x^\beta \tag{6.2}$$

$$\mathrm{d}s^2 = -\left(1 - 2\frac{U}{c^2}\right)(\mathrm{d}x^0)^2 + \left(1 + 2\frac{U}{c^2}\right)\left[(\mathrm{d}x^1)^2 + (\mathrm{d}x^2)^2 + (\mathrm{d}x^3)^2\right] \tag{6.3}$$

where ds is an infinitesimal space-time line element, τ is proper time as realised by an ideal clock, c is the velocity of light in vacuum, U is the sum of the gravitational potentials of the ensemble of masses and of a tidal potential generated by masses external to the ensemble, the latter potential vanishing at the barycentre. The four space-time coordinates are defined to be ($x^0 = ct$, x^1, x^2, x^3). The Einstein summation convention is used, implying summation over repeated indices. It should be noted that (6.3) gives only the first terms of a series, which is sufficient for the present level of observational accuracy. Higher order terms may be added as necessary. For time and frequency applications, this will be the case when clock stabilities reach some parts in 10^{19}.

6.2.2.2 Recommendation II

Recommendation II states that the space coordinate grid with its origin at the centre of mass of the Earth should show no global rotation with respect to a set of distant extragalactic objects, that the time coordinates for all coordinate systems should be derived from a time scale realised by atomic clocks operating on the Earth, and that the basic physical units of space-time are the second of the International System of units (SI) for proper time and the SI metre for proper length. This Recommendation should also apply to clocks on board terrestrial satellites.

6.2.2.3 Recommendation III

Recommendation III defines the scale units and origins of all time coordinates, and designates the solar system barycentric time coordinate and the geocentric time coordinate as Barycentric Coordinate Time, TCB, and Geocentric Coordinate Time, TCG, respectively. It should be noted that these time coordinates exhibit secular differences between themselves and with respect to TAI.

6.2.2.4 Recommendation IV

Recommendation IV defines Terrestrial Time (TT) a geocentric coordinate time scale differing from TCG by a constant rate, the scale unit of TT being chosen so that it agrees with the SI second on the rotating geoid. This constant rate is presently estimated to $6{,}9692904 \times 10^{-10}$ with an uncertainty of 1×10^{-17} ($1\ \sigma$).

As the theoretical time scales TCB, TCG, and TT are completely defined by the IAU Resolution A4, the path to realised time scales is immediate.

6.2.3 International Atomic Time

According to its definition, International Atomic Time (TAI) is simply a realisation of TT, apart from an offset of 32,184 s introduced for historical reasons. It is obtained by combining the data from an ensemble of about two hundred atomic clocks spread world-wide. To achieve this it is necessary to compare those clocks using a convention of coordinate synchronisation [ALLAN and Ashby, 1986]. This is defined as follows:

> *Two events fixed in some reference system by the values of their coordinates (t_1, x_1, y_1, z_1) and (t_2, x_2, y_2, z_2) are considered to be simultaneous with respect to this reference system if the values of the corresponding time coordinates are equal: $t_1 = t_2$. Two clocks are considered to be synchronised with respect to some reference frame if they simultaneously (in the above sense) produce the same time markers.*

In the vicinity of the Earth a geocentric, non-rotating reference frame, as defined in Resolution A4, Recommendation II of the IAU, is used for the synchronisation of clocks, and in particular for the calculation and dissemination of TAI.

To conform with the definition of TT, the scale unit of TAI is defined to be equal to the SI second as realised on the rotating geoid [BIPM, 1991]. To do this, data from the most accurate primary standards are individually corrected for the gravitational frequency shift arising from the elevation of the laboratory above the geoid and are then combined to form the scale unit of TAI.

6.2.4 Other coordinate time scales

Atomic coordinate time scales like TA(k), UTC, UTC(k), GPS time, GLONASS time, TT(BIPMxx)... *etc*. are time coordinates closely related to TAI and provided for different purposes. The scales TCG and TCB are related to TT and hence TAI by relativistic transformations [notes to Recommendations III and IV of the IAU Resolution A4, IAU 1992].

6.3 Generation of time scales

The practical problem is to generate a time scale from an ensemble of atomic clocks maintained in one or several laboratories. The efficient combination of the readings of the participating clocks requires [Tavella and Thomas, 1990a]:

- the definition of the expected qualities of the time scale,
- the characterisation of the available timing data,
- the design of an algorithm for data treatment.

6.3.1 Expected qualities

In general, the requirement is to generate a time scale which is as close as possible to an ideal time scale. The departure from an accumulation of ideal SI seconds on the rotating geoid can be estimated through computation of its "normalised frequency departure" at date t, commonly referred to as "frequency", defined as:

$$y(t) = \frac{\nu(t) - \nu_0}{\nu_0} \tag{6.4}$$

where $\nu_0 = 1$ Hz, and $\nu(t)$ is the reciprocal of the scale unit, expressed in SI second, of the time scale for date t.

Actual physical clocks have defaults that are minimised by combining their data to obtain a reliable, stable and accurate time scale. A separate, but important, point is the delay of access to the scale. For some purposes access must be immediate, for others a significant delay may be tolerable.

6.3.1.1 Reliability

Individual physical clocks may fail with immediate interruption of the time scales they deliver. Reliability thus calls for redundancy and eventually for national or international collaboration between laboratories maintaining atomic clocks.

The simplest solution to this problem is to replace the clock which has failed by another one. This is what is normally done in time laboratories which generate a UTC(k). Usually the UTC(k) is directly linked to the output of a physical clock, generally the best one of the ensemble on site, this being designated the "master clock". Its output is controlled by small, predetermined frequency and time steps, often through a microphase stepper, so as to steer UTC(k) to UTC. A change of master clock thus does not influence the output time scale if the microphase stepper is suitably programmed to handle the change.

More often, reliability is ensured by using an ensemble of clocks and computing an ensemble time. This time rarely has a physical realisation. In computing such an ensemble time, it is necessary to minimise the perturbations that result as clocks enter and leave the ensemble. Clearly the more clocks at the disposal of an ensemble, the less is the effect as one clock enters or leaves. For this reason there has been a general increase in the number of clocks in a given ensemble. For example, the number of clocks contributing to TAI was around 180 for some 10 years. Since beginning of 1993, however, the commercialisation of the new HP 5071A clock, from the firm Hewlett-Packard, has led to a steady increase in the number of clocks, which reached 237 in March 1994. A major consequence is a gain in reliability for TAI.

6.3.1.2 Stability

The stability of a time scale may be defined as its ability to maintain a constant scale interval, even if it differs from the ideal one. A measure of stability thus consists in the estimation of the dispersion of the frequency values $y(t)$ with time. Some statistical tools have been developed to estimate stability (see Chapter 4). They are efficient for the characterisation of the usual types of random noise which affect clock signals. The most common stability estimation tool is the two-sample, or Allan variance $\sigma_y^2(\tau)$ which depends upon the observational, or sampling, time τ.

The stability of an ensemble time scale depends upon the stabilities of the contributing clocks and the design of the algorithm used to generate it. The algorithm must, in particular, correctly handle any change in clock behaviour. The general considerations are detailed in 6.3.3, but the central idea is to generate a time scale more stable that any of the contributing clocks. This can be realised, but generally only for a given range of averaging times τ.

In principle, the concept of stability applies only to free-running time scales. A UTC(k) is, by definition, steered so it is affected by intentional frequency steps and its short- and middle-term stability is inevitably degraded. In addition, a crucial problem comes from the fact that time scale frequency values are always estimated or measured with respect to the frequency of another time scale or physical clock. The stability analysis of such comparative measures leads to evaluation of the coupled stability of the two time scales. Two cases may arise:

- The frequency of the time scale under test is evaluated by comparison with a time scale of better quality, such as those realised by primary frequency standards. The observed instability then may be ascribed totally to the time scale under test.

- The two time scales which are compared are supposed to be of similar quality. A technique for noise decoupling is then necessary. If the time scales involved in the comparison may be assumed to be completely independent, the *N*-cornered-hat technique [Barnes,1982; Allan, 1987] produces an estimate of the intrinsic stability of each element. If the independence is not verified, variances and covariances should be handled together for a complete analysis [Tavella and Premoli, 1994].

6.3.1.3 Accuracy

The accuracy of a time scale may be defined as its ability to maintain a mean scale interval as close as possible to its definition. For time scales which realise TT, the mean scale interval should be as close as possible to the SI second on the rotating geoid.

For primary frequency standards, accuracy is given by an uncertainty budget obtained through the evaluation of the physical effects that modify the output frequency with respect to the definition. When it is not possible to constitute such an uncertainty budget, accuracy is evaluated by comparison of the duration of the scale interval with the best realisation of the SI second provided by primary frequency standards. It is of course necessary to take into account the effect of the gravitational red shift on the primary standard frequency results in order to convert its realised SI second onto the geoid (null height). The accuracy of a time scale is generally given by a frequency difference between the time scale and the primary frequency standard, evaluated for averaging times corresponding to the best stability of the time scale, and by taking into account the uncertainty of the primary frequency standard.

Improvement in the accuracy of a time scale is generally carried out outside the main algorithm, which deals only with stability optimisation. This may be done by steering the frequency of the time scale on the frequency of a primary standard or of a more stable reference time scale. For this to be effective, the frequency corrections must be smaller than the frequency fluctuations of the time scale in order to avoid a degradation of its stability.

6.3.1.4 Delay of access

The delay of access to a time scale is linked to the quality of raw timing data and to the scientific purposes the time scale is supposed to fulfil.

Raw timing data is acquired according to a basic measurement cycle, whose duration ranges from several minutes to several hours, and is affected by measurement noise. Depending on the level of this noise, it may be necessary to smooth out raw measurements by accumulating data over several successive basic samples of measurement (see Section 6.3.2.). This delays access to the resulting time scale. In addition, it is useful to observe the behaviour of contributing clocks for a long period, both before and after the moment to which the data applies, in order to make the best use of their data. This also delays access.

What constitutes an acceptable delay of access to a time scale depends on its use. For a reference time scale, such as TAI, the requirement is extreme reliability and long-term stability. To match this purpose, the reference time scale relies on a large number of clocks of different types, located in different parts of the world. Data must therefore be collected and handled correctly, which takes time. The delay is thus considerable but is acceptable because of the ultimate quality obtained. For scientific studies inside a laboratory, however, it may be necessary to produce the time scale in near real time, immediately after the clock measurements, even if this impairs the long-term qualities of the scale.

6.3.2 Timing data

6.3.2.1 General form of timing data

Timing data takes the form of time differences between clocks. An atomic clock delivers a series of physical electric pulses separated from one another by a duration of 1 second, often designated as “series of 1 PPS”. Each pulse is an event with an associated number, a kind of label which is tied to it. This associated number is the reading of the clock for that particular event: for example, it may read as 1994 June 13 11 h 27 min 13 s. It can also be designated as the date of the event; its origin is arbitrary and is chosen to be convenient, but it is incremented by 1 second at each new pulse. Clock readings vary continuously and rapidly, so they can only be “caught in flight”. However, counters are available: they can be started with a given pulse coming from one clock and stopped with the pulse with the same label coming from another clock. A counter thus measures time differences which are proper quantities. These are thus measurable and expressed in SI units.

Denote $h_i(t)$ and $h_j(t)$ the time coordinates of the pulse labelled t, delivered by clock H_i, and of the pulse, with the same label, delivered by clock H_j, in a given reference frame. The interval of coordinate times:

$$x_{ij}(t) = h_j(t) - h_i(t) \tag{6.5}$$

is needed for the generation and dissemination of coordinate time scales.

At the present level of accuracy in clock comparisons, the coordinate quantity $x_{ij}(t)$ can be approximated using the interval of proper time issued from a counter, taking into account signal propagation delays for clocks separated by large distances [Petit and Wolf, 1994]; it thus can be expressed in SI units. In addition, in current practice, one does not specify a reference frame, and one designates the time coordinate $h_i(t)$ as the “reading of clock H_i at date t”, which, strictly, is not correct. For sake of conformity with the existing literature, we use the same designation in this text. However, the actual meaning of (6.5), which involves only coordinate quantities, should not be forgotten.

The quantities $x_{ij}(t)$ are the basic measurements used for time scale generation. They are obtained by time transfer methods, implemented between clocks located on the same site or in remote locations. Generally, a non-redundant network of time links is used, a given clock being compared only once with all the others at each date.

6.3.2.2 Comparison of clocks located on the same site

For computation of some of the TA(k) kept by national timing centres, all the contributing clocks are located on the same site. This is the case for the NIST (about 10 caesium clocks and 1 hydrogen maser), the SU (4 to 6 hydrogen masers), and the USNO (about 50 caesium clocks and 14 hydrogen masers). On each site, one clock is designated the master clock. Its output usually provides UTC(k), the local realisation of UTC. It serves also as a reference clock, to which the other clocks are compared in a star pattern as shown in Figure 6.1.

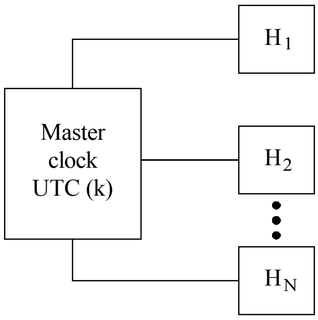


FIGURE 6.1

Comparison between clocks located on the same site

The measurements obtained take the form:

$$x_{ik}(t) = \mathrm{UTC(k)}(t) - h_i(t) \qquad \text{with} \quad i = 1, ..., N \tag{6.6}$$

at date t, N being the number of clocks.

The counters or time-interval meters normally used in timing laboratories provide measurements every second, or even more frequently, with accuracies ranging from 0.1 ps to 100 ps (1 σ) for one individual measurement (it has been shown that a 0.1 ps measurement noise can be obtained using the dual mixer time difference approach [Allan and Daams, 1975; Stein *et al.*, 1982]). If a time interval counter is used, averaging several readings, under the assumption that the residual measurement noise is white, may reduce the measurement noise to negligible levels. Such an averaging process is repeated with a basic measurement cycle τ_0 of order several hours; for example τ_0 = 2 h for the generation of AT1 at the NIST. However, it must be stressed that underlying temperature effects on time delays may negate some of the benefits of averaging.

6.3.2.3 Comparison of clocks located on remote sites

For the computation of some independent time scales TA(k), the contributing clocks are located in more than one laboratory. This is the case for the French TA(F), computed from 24 caesium clocks kept in 11 laboratories in France, for TA(CH) which includes data from 13 clocks kept in 3 Swiss laboratories, and for TAI computed from data reported by 45 national timing centres, maintaining between them about 230 atomic clocks [BIPM, 1993].

In addition to the pattern of Figure 6.1, used within the contributing laboratories, links, of a more elaborate nature, are required between distant UTC(k). This corresponds to Figure 6.2, and leads to measurements expressed in the form:

$$\begin{aligned} x_{ik_1}(t) &= \mathrm{UTC(k_1)}(t) - h_i(t) && \text{with } i = 1, ..., N_1 \\ x_{jk_2}(t) &= \mathrm{UTC(k_2)}(t) - h_j(t) && \text{with } j = 1, ..., N_2 \\ x_{k_1 k_2}(t) &= \mathrm{UTC(k_2)}(t) - \mathrm{UTC(k_1)}(t) && \end{aligned} \tag{6.7}$$

where t is the date, k_1 and k_2 the acronyms of the two laboratories being compared, and N_1 and N_2 are the number of clocks in each laboratory. Basic quantities $x_{ij}(t)$ defined in (6.5) are obtained by linear combination of the differences in (6.7).

There exist several methods for making distant time comparisons. Among the least accurate is that based on the reception of time signals emitted at radio frequencies, for example DCF77 emitted from Germany on

77.5 kHz [BIPM, 1993]. Terrestrial navigation signals such as the Loran-C were also widely used until about 1985. These had a precision on a single comparison of order 0.5 µs. In addition to this noise, huge seasonal variations were observed. Calibration of the equipment, receivers and emitters, was very difficult, and the accuracy obtained was characterised by an uncertainty (1 σ) of order several microseconds.

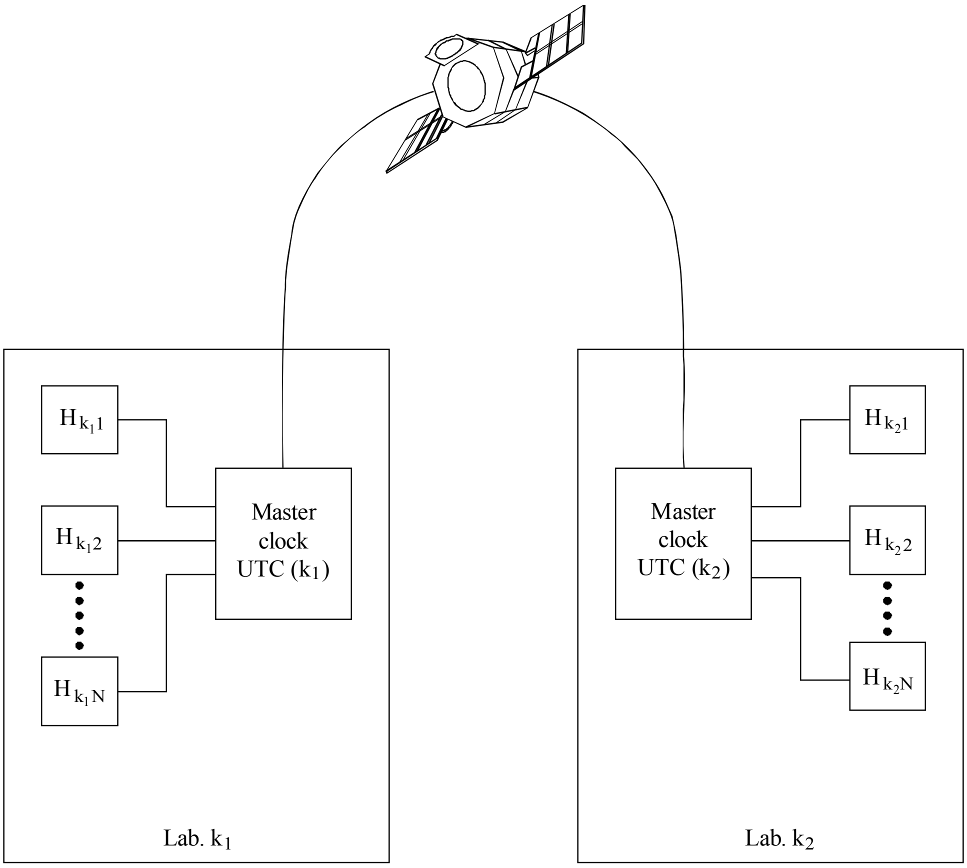


FIGURE 6.2

Comparison between clocks located on two different sites

The introduction of time transfer methods using satellite systems led to a major improvement in the precision, accuracy, and coverage of world-wide time metrology. All have a potential accuracy of order several ns (1 σ) or even better. At the beginning of 1994, four methods were in use or at an advanced planning stage:

- The Global Positioning System, GPS [Lewandowski and Thomas, 1991]. The GPS is an American military navigation system based on satellite ranging using on-board atomic clocks. Since GPS was declared operational, in December 1993, it has been able to provide position, velocity and time instantaneously and continuously anywhere on or above the Earth. In particular, the observation of any GPS satellite gives access to the time scale known as GPS time, which is closely steered to UTC(USNO). For timing applications, the GPS is used according to the common-view method [Allan and Weiss, 1980], which makes it possible to overcome partly the intentional degradation brought to the satellite signals. In 1994, this method was routinely used by most of the national timing laboratories of the world, leading to uncertainties on time comparisons of the order of several ns.

- The GLObal NAvigation Satellite System, GLONASS [Daly *Et Al.*, 1992; Lewandowski *et al.,* 1993]. The GLONASS is the Russian equivalent to GPS but has no intentional signal degradation. Commercial time receivers are not yet available, so the system is not used widely.

- Two-Way Satellite Time Transfer via a geostationary satellite, TWSTT [Kirchner *et al.*. 1991;De Jong, 1993]. The TWSTT requires, on site, a station for emission and reception of microwave signals in the telecommunication band, and a satellite channel for signal repetition on board.

- LAser Synchronisation from Satellite Orbit, LASSO [Baumont *et al*., 1993]. The LASSO requires, on site, a laser shooting station and a satellite equipped with stable oscillators, counters, and light retro-reflectors.

6.3.2.4 Smoothing of data measurement noise

Data of comparison between remote clocks exhibits a measurement noise whose origin is the time transfer method. It is necessary to remove this noise in order:

- to take advantage of the full quality of the clocks being compared, and

- to avoid the injection of measurement noise into the time scale itself, which would degrade its short-term stability.

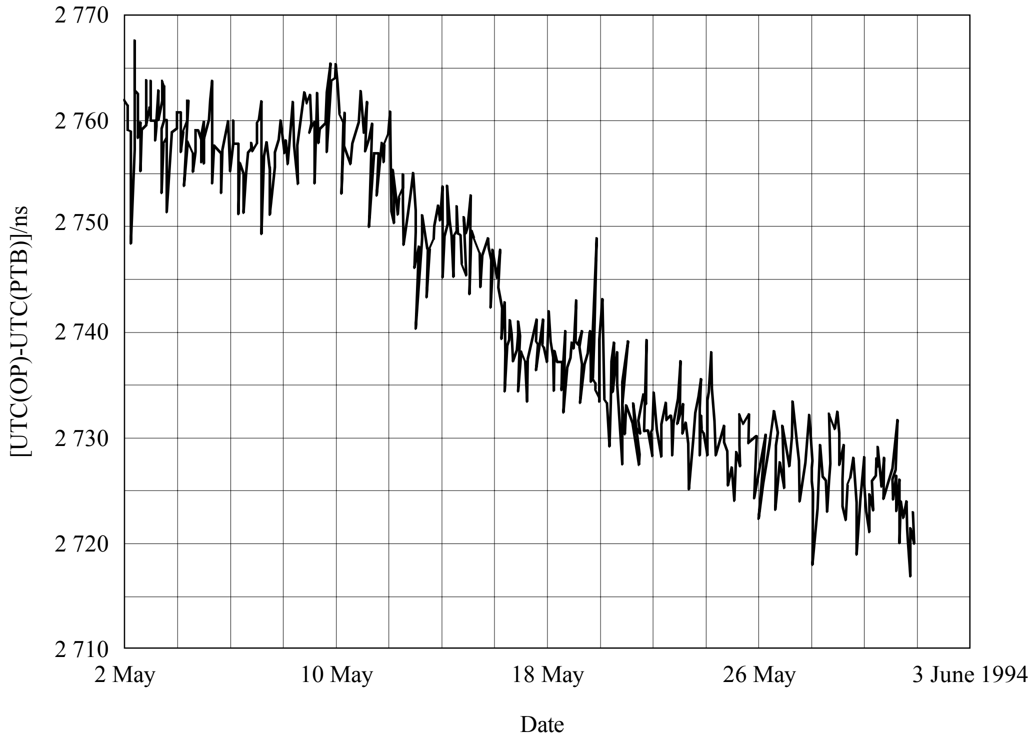
Efficient smoothing of measurement noise requires its statistical analysis. Here we illustrate this smoothing for the frequent case of GPS common-view time transfer between two laboratories. The example chosen here corresponds to the European time link between the OP, Paris, France, and the PTB, Braunschweig, Germany. Figure 6.3.a shows the raw common-view values, obtained for a thirty-day period in May 1994. They correspond to about 24 daily tracks of the international GPS common-view schedule No 22. This data is first treated by the computation of the Allan standard deviation (see Chapter IV), with the hypothesis of equally-spaced data, distant by τ_0 = 1/24 day. In the log-log plot of Figure 6.3.b, the Allan standard deviation values $\sigma_y(\tau)$ lie on a straight line of slope –1 for averaging times $\tau_0 \leq \tau \leq 1$ day. This indicates the presence of phase noise for τ smaller than one day. The actual performance of the master clocks in the OP and the PTB is not dominated by phase noise for such averaging times, and thus becomes accessible as soon as the phase noise, whose origin is the time comparison method, is smoothed out. For this purpose it is sufficient to take average values on consecutive raw data $x_{\alpha OP\,PTB}(t)$ covering one day. This leads to time link values $x_{OP\,PTB}(t')$ reported at the dates t' corresponding to the middle of successive days, with smoothed random noise arising from the GPS common-view method. The level of white phase noise is estimated from the modified Allan standard deviation $Mod.\sigma_y(\tau_0)$, using (see Chapter 4):

$$\sigma_x = \tau_0 \frac{\mathrm{mod}\,\sigma_y(\tau_0)}{\sqrt{3}} \tag{6.8}$$

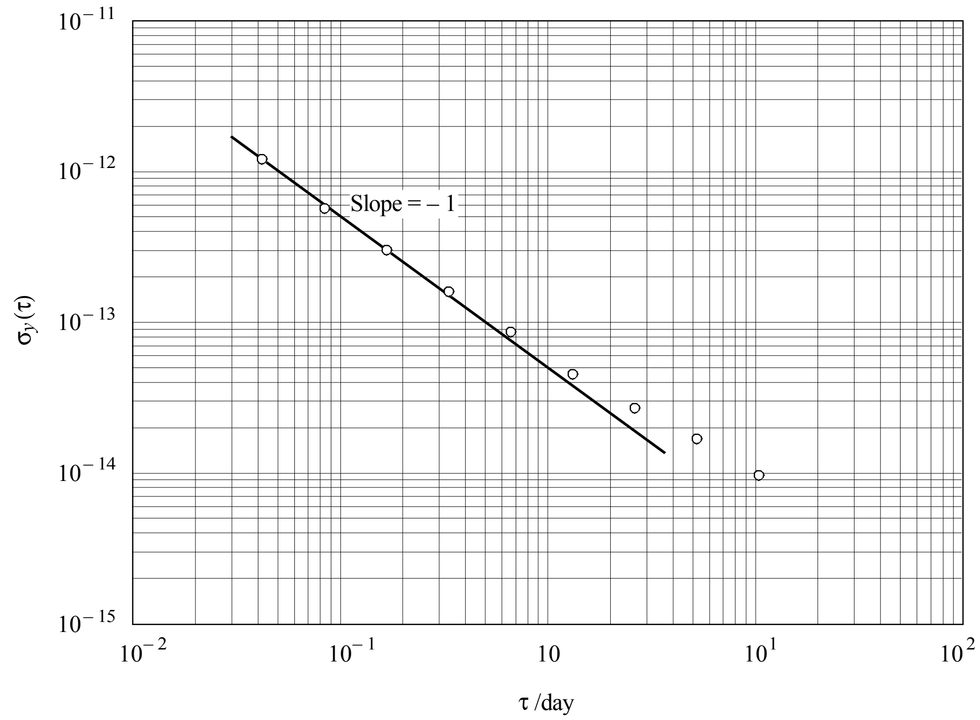
which gives σ_x = 2.6 ns for this example. The residual white phase noise on daily average would be of order $\sigma_x/\sqrt{24}$, which is below 1 ns, if the 24 observations were giving independent estimates of the time difference, which is not totally the case. Anyway, the residual noise, after averaging, is negligible compared with the daily clock performance.

The above example shows that, although GPS time data are taken with a rather short measurement cycle τ_0, equal to 1 hour in this particular case, measurements of interest, *i.e.* ones actually representing the quality of the clocks themselves, are available only with a basic period T_0 of order 1 day. For long-distance GPS links, T_0 ranges from 2 to 3 days in the best cases, when using measured ionospheric delays and precise satellite ephemerides. Before the introduction of GPS, durations T_0 as long as 50 days were necessary to reduce the Loran-C measurement noise to acceptable levels.

In all cases, the measurement noise affecting timing data is smoothed out before applying the main algorithm for computation of the time scale.



a) Raw timing data for month May 1994



b) Corresponding Allan deviation

FIGURE 6.3

Common-view GPS time transfer between the OP and the PTB

6.3.3 Stability algorithm

Suppose an ensemble of N clocks: at date t, corresponding timing data are the (N-1) measurements $x_{ij}(t)$, with respect to clock H_j, chosen to be non-redundant, and given by (6.5):

$$x_{ij}(t) = h_j(t) - h_i(t)\ ,\ i = 1, ..., N,\ i \neq j$$

Suppose TA is the resulting software time scale; it should be computed for date t from the optimal combination of the $x_{ij}(t)$. The N time differences:

$$x_i(t) = \mathrm{TA}(t) - h_i(t)\ ,\ i = 1, ..., N \qquad (6.9)$$

give access to TA at date t. The $x_i(t)$ are the unknowns.

Suppose TA is known for a given date t_0 where the measurements $x_{ij}(t_0)$ were available and treated. Measurements $x_{ij}(t)$ are now taken for a following date t, $t > t_0$. The dates t and t_0 are usually separated by a duration T greater than T_0. The problem is to design an algorithm, able to handle the timing data $x_{ij}(t)$ for the generation of TA at date t.

A time-scale algorithm is generally designed to ensure the best stability of the time scale, accuracy being treated externally as explained in 6.3.1.3. It is important to stress that there is no general best solution in the design of time scale algorithms. Rather, good design represents a series of choices matched to the purpose for which the time scale is to be used. An algorithm designed to provide a time-reference standard is unlikely to satisfy the requirements of those whose interest is the provision of a service for research. One critical choice is, for example, whether the algorithm must supply the time scale in real-time, or close to it, or whether a delayed scale is acceptable. In all cases, however, the statistical treatment of clock data requires at least [Tavella and Thomas, 1991a]:

- the definition of an average time scale,
- the choice of a duration between two updatings of the time scale,
- the specification of a procedure to optimise the contribution of each clock, and
- the implementation of a filter on each clock frequency to provide a means of frequency prediction, with appropriate compensation if frequency drift is present.

The time scale algorithms which are in use in timing centres rely upon two basic assumptions:

- Measurement results $x_{ij}(t)$, given in (6.5), are affected by intrinsic noise which is negligible with respect to the clock noise;
- Clocks are independent and the corresponding data series are uncorrelated. This assumption is conceptually true as each clock is an independent box in which atoms run and "lock" the frequency generated inside. But, in 1989, the Comité Consultatif pour la Définition de la Seconde recommended that a study be made of possible correlations among clocks. Through a survey on the behaviour of the clocks contributing to TAI [Tavella and Thomas, 1990b, 1991b], some correlations between clock frequencies were detected. These correspond mainly to responses to changes in the environmental conditions experienced by the clocks. Since several years, efforts have been pursued to improve clock independence either through better control of the environment or through the realisation of less sensitive atomic clocks [De Marchi, 1988].

In the following, we refer to examples for which extensive documentation can be found. These are, in particular, the algorithm ALGOS(BIPM) [Guinot and Thomas, 1988, Tavella and Thomas, 1991a], which produces the international reference TAI at the BIPM, and the algorithm AT1(NIST) [Varnum *Et Al.*, 1987; Weiss *et al.*, 1989], which produces the real-time time scale AT1 at the NIST. The ALGOS(BIPM) treats data from a large number of clocks spread world-wide. It is designed for extreme long-term stability and a delay of access of several weeks is acceptable for TAI delivery. The AT1(NIST) treats data from about 10 clocks kept on the same site. It is designed for scientific experiments requiring real-time access to AT1.

6.3.3.1 Definition of an average time scale

To match the definitions of time scales given in the introduction and in section 6.2, the reading of an atomic time scale TA may be theoretically written, at date t, as the weighted average of the readings of the contributing clocks:

$$\mathrm{TA}(t) = \sum_{i=1}^{N} \omega_i(t) h_i(t) \tag{6.10}$$

The basic mathematical definition given in (6.10) plays a fundamental role in the development about time scale algorithms which is detailed in the following. Some ensemble times, such as the old TA(NIST) [Jones and Tryon, 1983, 1987], stopped in 1993, or GPS time [Feese *et al.*, 1991], do not make use of a similar average definition. These cases are not considered here.

Relative weights $\omega_i(t)$, i = 1, ..., N, are introduced in order to discriminate between clocks according to their intrinsic qualities. They satisfy the following relation:

$$\sum_{i=1}^{N} \omega_i(t) = 1 \tag{6.11}$$

The time of a clock is in general in error from some defined time scale because of both systematic deviations as well as random deviations. The weighting function in (6.10) is typically chosen to optimise stability and is not dependent on the systematic deviations (e.g. frequency offset, frequency drift) but only on the random deviations. Hence, if the weight of a clock is changed or if a clock is added or subtracted, the resulting time scale computation will be adversely affected, since the weighting function in (6.10) apply to systematic deviations as well [Allan *Et Al.*, 1974; Guinot, 1987]. It follows that (6.10) is not satisfactory for actual clock ensembles, for which changes of weights are unavoidable. Equation (6.10) must thus be completed as follows:

$$\mathrm{TA}(t) = \sum_{i=1}^{N} \omega_i(t) \left[h_i(t) + h_i'(t) \right] \tag{6.12}$$

where $h_i'(t)$ is a time correction added at date t to the reading of clock H_i, and designed to ensure time and frequency continuity of TA at a previous date t_0 when weights have been changed [Guinot and Thomas, 1988]. The correction $h_i'(t)$ is written as:

$$h_i'(t) = x_i(t_0) + y_{ip}(t) \cdot (t - t_0) \tag{6.13}$$

where $x_i(t_0) = TA(t_0) - h_i(t_0)$ is known, since it results from the computation of TA at date (t_0, and where $y_{ip}(t)$ is the predicted frequency of clock H_i, relative to TA, over the interval [t_0, t]. The frequency $y_i(t)$ of clock H_i, relative to TA, over the interval [t_0, t] can be estimated from:

$$y_i(t) = \frac{\left[\mathrm{TA}(t) - h_i(t)\right] - \left[\mathrm{TA}(t_0) - h_i(t_0)\right]}{t - t_0} \tag{6.14}$$

Until TA is computed at date t, $y_i(t)$ is unknown. It is thus necessary to predict it according to the past behaviour of clock H_i. This predicted frequency is denoted $y_{ip}(t)$ and appears in (6.13).

Equations (6.5) and (6.9) and (6.12) lead to the following system of equations, assuming no measurement noise:

$$\sum_{i=1}^{N} \omega_i(t) x_i(t) = \sum_{i=1}^{N} \omega_i(t) x_i(t_0) + \sum_{i=1}^{N} \omega_i(t) y_{ip}(t) \cdot (t - t_0) \tag{6.15}$$

$$x_i(t) - x_j(t) = x_{ij}(t)$$

System (6.15) is deterministic with N equations and N unknowns. The solution is unique and the results are the time differences $x_i(t)$, $i = 1, ..., N$, which give access to TA for date t. The difference between clock H_j and TA is explicitly given by:

$$x_j(t) = \sum_{i=1}^{N} \omega_i(t)\left[h_i'(t) - x_{ij}(t)\right] \tag{6.16}$$

System (6.15) can be found in most of the algorithms used around the world, for instance in the algorithms used for computing AT1 at the NIST [Varnum *et al.*, 1987; Tavella and Thomas, 1991a], TA(F) [Granveaud, 1986] at the OP, TAI [Guinot and Thomas, 1988] at the BIPM, TA(AUS) at the ORR [Luck, 1979], TA(CRL) at the CRL [Yoshimura, 1980], and A.1(MEAN) at the USNO [Percival, 1978].

The advantages of optimising both the short-term and the long-term stability have been demonstrated [ALLAN *et al.*, 1974], and new developments of time-scale algorithms have been recently proposed, which address the possibility of using short-term and long-term weighting procedures [Wei Gu, 1992, Stein 1992]. At the USNO a new algorithm is used in which the ensemble time is re-evaluated every hour from the past 75 days: the weights are modified according to a quadratic variation with time in order to match the short-term and long-term qualities of different clock types (caesium standards and hydrogen masers). The update for the last hour is used to steer the master clock [Breakiron, 1991].

For some algorithms the definition of the time scale is used together with specific filters which act on the raw timing data, which were not smoothed before. This is the case of KAS-1 [Stein, 1988; Stein *et al.*, 1989] for which a Kalman filter is implemented. In other cases the Kalman formalism is used for the resolution of (6.16) in order to improve the stability of the time scale, as is done in KAS-2 [Stein,1992], or to evaluate uncertainty of the estimates and detect abnormal behaviours, as is done in TA2(NIST) [Weiss and Weissert, 1994]. In the following we restrict our discussion to the “classical” and well established ensemble algorithms which rely on (6.15), referring the interested reader to the literature for further developments.

Since the definition of the time scale, and thus the resulting system of equations, nearly always takes the same form, the specificity of a given algorithm lies in the choices made for:

- the length of the time interval $[t_0, t]$,
- the weights attributed to clocks,
- the way clock frequencies are predicted, and the way frequency drift is dealt with.

These choices are closely related to the purposes for which the time scale is designed.

6.3.3.2 Length of the basic interval of computation

In previous sections, two basic durations have already been defined:

- τ_0, duration of the basic measurement cycle.
- T_0, minimum duration over which raw data should be averaged in order to smooth out the measurement noise sufficiently to reach the intrinsic qualities of the clocks being compared. Orders of magnitude for T_0 are several minutes to several hours inside a laboratory, 12 hours to 1 day between two laboratories linked via short-distance GPS common views, and several days between two laboratories linked via long-distance GPS common views.

Also defined are two dates:

- t_0, date for which TA is known.
- t, a following date ($t > t_0$) for which smoothed timing measurements are available and for which TA is to be computed by solving (6.15).

The update interval $T = t - t_0$ is, in general, of the same order of magnitude and slightly longer than T_0. Its length is thus directly linked to the quality of timing data. It is for example:

- $T = 2$ hours for AT1(NIST), which only uses timing data taken on site,

- $T = 1$ day for TA(F), which uses timing data from all over France, the maximum baseline between laboratories being of order 1 000 km,

- $T = 10$ days for TAI, which uses timing data from all over the world, the maximum baseline between laboratories being of order 6 000 km.

Another requirement is efficient characterisation of the behaviour of the participating clocks in order to weight them correctly and efficiently predict their frequencies relative to TA (see the following sections). It is thus often necessary to observe clocks over a duration longer than T. One then has two possibilities as follows.

Consider an integer n greater than 1.

6.3.3.2.1 Update of TA every interval of duration *T*

A memory of the last n intervals of duration T is retained. The time scale is delivered in near real-time, with a delay no greater than T, but it is based only on the past behaviour of contributing clocks. There is no re-processing and no post-processing. The weights and the frequency prediction are valid for an interval of duration T. The resulting algorithm is thus dynamic and adaptive at intervals of T.

The advantage of this approach is that the time scale is accessible in real time. The disadvantage is that it is not possible to take into account the abnormal behaviour of one clock before it registers on the scale. A stable clock which suddenly presents a frequency jump can thus sweep along the time scale before the anomaly is detected.

This approach is used for AT1, for which $T = 2$ hours and $nT \approx 10$ days ($n \approx 120$). The problem of detecting abnormal behaviour is partly solved in an updated algorithm AT2, conceived and tested at the NIST [Weiss and Weissert, 1991].

6.3.3.2.2 Update of TA when the interval of duration *nT* is ended

The $(n + 1)$ dates included in the interval are treated as a whole. This delivers a deferred-time time scale, computed in post-processing. The weight and the frequency prediction of a given clock are valid for an interval of duration nT. They are changed for the following interval of duration nT, but are equal for all dates included in a given interval of duration nT. The clock behaviour observed during the whole interval of computation is taken into account. The resulting algorithm is dynamic and adaptive *a posteriori* at intervals of duration nT.

The advantage is the possibility of taking into account any abnormal behaviour of clocks occurring during this period. The disadvantage is the access to the time scale in deferred-time for the $(n + 1)$ dates included in the interval of computation.

This is the case for TA(F), for which $T = 1$ day and $nT = 30$ days ($n = 30$). The TAI is computed by the same process with $T = 10$ days and $nT = 60$ days ($n = 6$).

Another algorithm, at the NIST, uses both of these processes. This is TA2 which relies on the algorithm AT2 (AT1 plus abnormal behaviour detection), operating with $T = 2$ hours and $nT \approx 10$ days, run forward and backward over a duration of one month [Weiss and Weissert, 1994]. It follows that there is an iterative re-processing of the data throughout the month. The NIST thus has at its disposal two time scales, the real-time AT1 and the deferred-time TA2, the computation of the previous TA(NIST), based on a Kalman filter [Barnes, 1982], being discontinued since mid-1993.

Most of the algorithms used in national laboratories adopt the first choice, updating TA(k) in real time or in near real-time without post-processing. In addition, some algorithms, as those for TAI at the BIPM, A.1(MEAN) at the USNO, or TA2 at the NIST, choose an iterative procedure to evaluate weights and frequency predictions: this takes the form of successive recomputations of TA for the same interval, with detection of outliers at each step, until results converge [Tavella and Thomas, 1991a].

6.3.3.3 Weighting procedure

6.3.3.3.1 General ideas

Since time scale algorithms are designed to optimise frequency stability, each clock should be weighted according to its own frequency stability. The weight attributed to a given clock is thus basically chosen to be inversely proportional to its frequency variance σ_i^2.

$$\omega_i = \frac{1/\sigma_i^2}{\sum_{k=1}^{N} 1/\sigma_k^2}, \; i = 1, ..., N \tag{6.17}$$

The reason for this is that, if the contributing clocks are independent and if weights are not artificially limited, the frequency variance of the resulting time scale may be written as:

$$\frac{1}{\sigma_{TA}^2} = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} \tag{6.18}$$

which means that the time scale is, in principle, more stable than any contributing element. The choice of the variance type (classical, filtered, or Allan) depends on the purposes for which the time scale is generated, and can thus differ according to the algorithm which is considered. However, there are two limiting factors as follows.

The frequencies of clock H_i, used for the computation of its frequency variance, are estimated over an interval of duration τ. According to (6.18), the stability of the resulting time scale is optimised for averaging times close to τ. It is thus of pre-eminent importance to determine for which τ values the contributing clocks present their best stabilities, and to define what objective of stability the time scale should fulfil. In other words, the optimisation of both short-term and long-term stability could call for contributions from different types of clocks, treated according to different procedures in the algorithm. This is the case for the software UTC(USNO) computed at the USNO [Breakiron, 1991], and also for the KAS algorithms [Stein, 1992].

The frequencies of clock H_i, used for the computation of its frequency variance, are estimated by comparison with a reference. Very often this reference is the time scale itself, because its stability is supposed to be better than that of the contributing clocks. It follows that the computed variance is inherently biased [Yoshimura, 1980] and ceases to represent the true quality of the clock. This is the so-called “clock-ensemble correlation” effect. An approach to the derivation of this effect has been published [Tavella *et al.*, 1991], and gives:

$$\sigma_{i,bias}^2 = \sigma_{i,true}^2 (1 - \omega_i) \tag{6.19}$$

where $\sigma_{i,bias}^2$ and $\sigma_{i,true}^2$ are the “biased” and "true" frequency variances of clock H_i.

The effect of clock-ensemble correlation is proportional to the relative contribution of the clock within the ensemble. If not taken into account, a very stable clock is progressively more heavily weighted, which threatens the reliability of the time scale. The correction factor of (6.19) appears in most algorithms used in national laboratories, sometimes with a multiplication factor close to 1 [Tavella and Thomas, 1991a]. However, it does not intervene in the TAI algorithm because the number of contributing clocks and the implementation of an upper limit of weight lead to a maximum contribution, ω_i, of a given clock, which has been smaller than 1% since beginning of 1993, and is thus negligible with respect to 1.

In addition to the fundamental aspects that have just been discussed, the weighting procedure must obey some other rules. The most important is the implementation of an upper limit of weight, necessary practically to make the time scale rely on the best clocks and yet avoid giving a predominant role to any one of them. Another is an objective criterion to safeguard the time scale against the possible abnormal behaviour of some clocks. It is important to stress that the existence of an upper limit of weight safeguards reliability but invalidates (6.18). It can thus lead to a time scale TA that is no better than the best single contributing clock.

To illustrate, we take the examples of the algorithms AT1(NIST) and ALGOS(BIPM) for which a complete comparison is available [Tavella and Thomas, 1991a].

6.3.3.3.2 Weighting procedure in AT1 (NIST)

In AT1(NIST), the weights used for the computation of AT1 at date t are deduced from the results of the computation of AT1 at date t_0 ($t - t_0 = T$). The weight $\omega_i(t)$ of clock H_i is obtained from (21) where $\sigma_i^2(t)$ results from an exponential filter written as:

$$\sigma_i^2(t) = \frac{1}{A+1}\left[\delta_i^2 + A \cdot \sigma_i^2(t_0)\right] \tag{6.20}$$

with:

$$\delta_i = \left|(t_0) - y_{ip}(t_0)\right| + {K_i}/{T} \tag{6.21}$$

The exponential filter is used to de-weight the past behaviour of the clock. Its time constant A is usually set to 20 to 30 days. The term δ_i contains the shift between the actual frequency of clock H_i and its predicted value, thus giving an estimation of the predictability of the clock over T. The term K_i, added in (6.21), takes into account the correlation between the ensemble time and clock H_i. This is absolutely necessary in the AT1(NIST) algorithm, which is designed for the treatment of a small number of clocks (≈ 10) and where the maximum contribution of a given clock can reach 20%. Since very recently, the term K_i has been chosen according to (6.19) in both the AT1 and TA2 algorithms [Weiss and Weissert, 1994].

The AT1(NIST) weight determination keeps no memory of the absolute values of past frequencies, but rather relies on frequency variations. This is similar to the difference between an Allan variance and a classical variance. Although the clock frequency instability is tested, it should be noted that some information about long-term systematic variations could be lost.

The use of an exponential filter for the weight determination is efficient because it de-weights the past: if a clock has a frequency "accident", and thus is intentionally de-weighted, its de-weighting is progressively removed over an interval of several integration times. In AT2(NIST) and TA2(NIST) a frequency-step detection is explicitly introduced [Weiss and Weissert, 1994]: the basic idea is to detect a frequency difference larger than 4 times the level of frequency noise observed for that clock. In addition, an upper limit of weight is introduced in AT1(NIST) for the sake of reliability.

6.3.3.3.3 Weighting procedure in ALGOS (BIPM)

As already noted, ALGOS(BIPM) operates in post-processing, treating as a whole measurements taken over a basic period $nT = 60$ days. Measurements are available every $T = 10$ days, on the MJD ending with 9. The time scale is updated for each of the six dates t included in the two-month period under consideration: $t = t_0 + mT$, with $m = 1, 2, 3, 4, 5, 6$. The date t_0 is the last date of the previous two-month interval, for which the time scale is maintained and not updated. The separation between updates is thus 10 days, but the gap between calculations is 60 days.

In ALGOS(BIPM), the weight $\omega_i(t)$ of clock H_i is constant over the two-month interval I of computation: it is thus valid for the seven dates $t = t_0 + mT$, with $m = 0, 1, 2, 3, 4, 5, 6$, continuity at t_0 being ensured by the clock frequency prediction. It can be written $\omega_i(I)$ and obeys (6.17), where $\sigma_i^2(I)$ are individual classical variances computed from six consecutive two-month frequencies of the clock H_i. These are the frequencies computed over interval I and the five previous two-month intervals. As the frequency over the interval I is not yet known, an iterative process is used [Tavella and Thomas, 1991a] which begins with the weights obtained at the previous two-month computation, ending at date t_0; this gives an indication of the behaviour of each clock during the interval I and thus makes it possible to refine weights in the following iterations.

The ALGOS(BIPM) weight determination uses clock measurements covering a full year, so annual frequency variations and long-term drifts can lead to de-weighting. This has helped to reduce the seasonal variation of TAI observed during the seventies and the eighties. In addition, the choice of 60 days, initially chosen to smooth the Loran-C data, corresponds to a good averaging time for the detection of frequency anomalies. Sampling over 60 days thus allows optimisation of TAI stability in the long-term. With the increased use of GPS common-view links and of the newly designed HP clocks, *nT* could be reduced to 30 days. The weight could then be determined with 12 one-month samples.

In ALGOS(BIPM), the term for the clock-ensemble correlation of (6.19) is negligible and is thus not introduced [Tavella *et al.*, 1991]. There is an upper limit of weights which corresponds to a minimum variance $\sigma_i^2(\mathrm{I})$ of 3.66×10^{-14}, which could be changed, if called for by improvements in clock performance. An algorithm to detect abnormal behaviour is also implemented: this tests frequency changes [Tavella and Thomas, 1991a].

To conclude, the weights used in AT1(NIST) and ALGOS(BIPM) obey the same rules, in particular: optimisation of the stability, detection of abnormal behaviour, minimisation of the clock-ensemble correlation. The specific choices which have been made, match the available timing data and fulfil the fundamental requirement of access to a time scale in real time or deferred time.

6.3.3.4 Frequency prediction

6.3.3.4.1 General ideas

The way the frequency of clock H_i is predicted depends on its statistical characteristics and on the duration for which the prediction should be valid. There are several pure cases:

- The predominant noise is white frequency noise: this is the case for commercial caesium clocks for averaging times τ ranging from 1 day to 10 days. The most probable frequency, estimated over an interval of duration τ, for the following τ interval is then given by the mean of the frequency values observed over a number of previous intervals of duration τ.

- The predominant noise is random walk frequency modulation: this is the case for most commercial caesium clocks for averaging times τ ranging from 20 days to 70 days. The most probable frequency value for the following τ interval is then the last frequency value estimated over the previous interval of duration τ.

- The predominant frequency deviation is a linear drift: this is the case for some hydrogen masers for averaging times τ longer than several days. The most probable frequency for the following τ interval is then the last frequency calculated over the previous interval of duration τ corrected by a term deduced from the estimated frequency drift.

To optimise an ensemble it is thus necessary to have a good knowledge of the behaviour of the contributing clocks, and to be astute in selecting suitable modes of frequency prediction for the different clock types.

To illustrate, we consider the algorithms AT1(NIST) and ALGOS(BIPM). A comparison is also available in [Tavella *et al.*, 1991a].

6.3.3.4.2 Frequency prediction in AT1 (NIST)

For AT1(NIST), the predicted frequency $y_{ip}(t)$ of clock H_i, for computation of AT1 at date t, is deduced from the results of the computation of AT1 at date t_0, with $t - t_0 = T$. This is obtained from an exponential filter written as:

$$y_{ip}(t) = \frac{1}{B_i + 1}\left[y_i(t) + B_i \cdot y_{ip}(t_0)\right] \qquad (6.22)$$

The predicted frequency of clock H_i is an average of the frequencies of clock H_i over past periods with an exponential weighting. The time constant B_i of the exponential filter depends on the statistical properties of clock H_i and may thus differ from one clock to another. It allows an optimal estimation of the long-term behaviour of the clock, since it corresponds to the averaging time for which the clock reaches its flicker floor or for which a good estimation of the random walk component is possible.

6.3.3.4.3 Frequency prediction in ALGOS (BIPM)

As already noted, ALGOS(BIPM) operates in post-processing, treating as a whole measurements taken over a basic period $nT = 60$ days. As with its weight, the predicted frequency of clock H_i is constant over the two-month interval I of computation; it is thus valid for the seven dates $t = t_0 + mT$, with $m = 0, 1, 2, 3, 4, 5, 6$, and can be written $y_{ip}(I)$.

In ALGOS(BIPM), the predicted frequency used for the present two-month interval is equal to the frequency obtained over the previous two-month interval as a one-step linear prediction. This is the optimal prediction for averaging times of two months, for which the predominant noise is random walk frequency modulation. All clocks contributing to TAI are subjected to the same mode of frequency prediction; however, changes in the procedure are under discussion, in particular the introduction of a frequency drift estimation to predict the frequencies of hydrogen masers.

To conclude, the modes of frequency prediction in AT1(NIST) and ALGOS(BIPM) differ because each is adapted to the length of its own basic interval of computation and so to the statistical properties of the clocks over such averaging times.

6.3.4 Accuracy of the scale interval of a time scale

Improvement of the accuracy of a time scale is generally carried out outside the main algorithm, which deals only with optimising stability.

For TAI, it is carried out by frequency steering the free-running time scale derived from the stability algorithm ALGOS(BIPM). The frequency corrections are smaller than the frequency fluctuations of the time scale to avoid a degradation of its stability. They are decided after comparison of the frequency of the computed time scale with a combination of the frequencies of primary frequency standards, continuously operating or occasionally evaluated, all around the world [Azoubib *et al.*, 1977]. In this exercise, the effect of the gravitational red shift on the primary standard frequencies is taken into account. Only one frequency steering correction was applied in 1993: it amounted to 0.5×10^{-15}. The accuracy of TAI is expressed in terms of the mean duration of its scale unit, computed for two-month intervals, in SI seconds on the rotating geoid. It is published in the successive volumes of the *Annual Report of the BIPM Time Section*. For example, the mean duration of the TAI scale unit was equal to $(1 + 0.2 \times 10^{-14})$ SI second on the rotating geoid for the interval May-June 1993, with an uncertainty (1 σ) equal to 1.3×10^{-14}.

For the NIST atomic time scales, accuracy is ensured by comparisons with the primary frequency standards NBS-6 and NIST-7.

6.3.5 Examples

6.3.5.1 Stability of some independent time scales

Allan standard deviations have been computed using the time comparison values between TAI and, respectively, TA(F), AT1, TA(PTB), and A.1(MEAN), collected in [BIPM, 1993].

The TA(F) is computed from 23 caesium clocks in laboratories distributed all over France, with an algorithm similar to ALGOS(BIPM). The minimal value of the Allan standard deviation is:

$$\sigma_y(\tau \approx 40 \text{ days}) \approx 8 \times 10^{-15} \tag{6.23}$$

The AT1 is computed from about 10 caesium clocks maintained on a single site, using the AT1(NIST) algorithm. The minimal value of the Allan standard deviation is:

$$\sigma_y(20 \text{ days} \leq \tau \leq 40 \text{ days}) \approx 5 \times 10^{-15} \tag{6.24}$$

The TA(PTB) is not derived from a time scale algorithm. It is simply the output of the primary frequency standard PTB CS2, which operates continuously as a clock. The minimal value of the Allan standard deviation is:

$$\sigma_y(80 \text{ days} \leq \tau) \approx 6 \times 10^{-15} \tag{6.25}$$

The TA(USNO) is the time scale A.1(MEAN) computed from about 50 caesium clocks (36 of them are HP 5071A units) and 14 hydrogen masers kept on site, with an algorithm which uses a double weighting procedure for optimisation of both short-term and long-term stability. The minimal value of the Allan standard deviation is:

$$\sigma_y(\tau \approx 80 \text{ days}) \approx 5 \times 10^{-15} \tag{6.26}$$

Since the values of Allan standard deviations given here describe the time differences between TAI and the independent time scales, the part of instability coming from TAI is not separated from that coming from the individual TAs. Application of the N-cornered-hat technique allows this separation provided that the time scales entering in the computation are statistically independent. Figure 6.4 shows the Allan standard deviation values for TAI obtained with a 4-cornered-hat technique, using data from comparisons between TAI and AT1, TAI and TA(SU) and TAI and TA(PTB) for the period January 1993 – April 1994. The values obtained are always smaller than 6×10^{-15}.

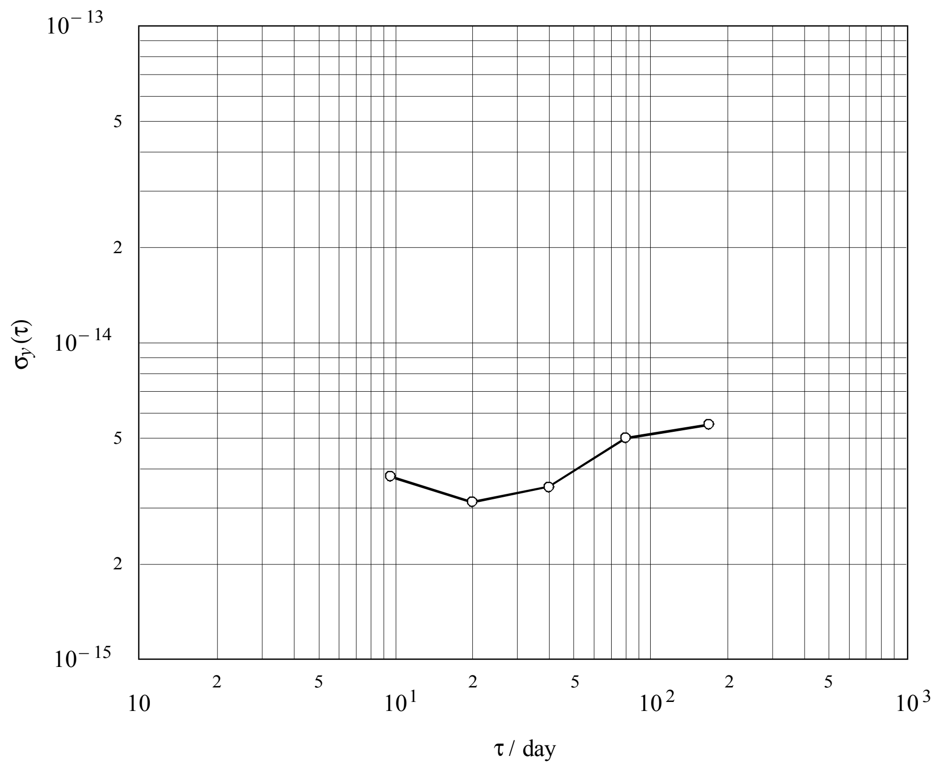


FIGURE 6.4

Allan standard deviation values for TAI

It is important to note that the reported values from the Allan deviation have decreased considerably for most of the independent time scales in the past few years. For TAI, since the introduction of the new HP 5071A clocks and the use of active auto-tuned hydrogen masers, the values have also fallen substantially. It remains that hydrogen masers, though presenting outstanding short-term and intermediate- term stabilities, bring some problematic long-term frequency drift to TAI.

6.3.5.2 Steering of some local representations of UTC

Figure 6.5 show two examples of time variations of comparisons between UTC and UTC(k) over one year ending in May 1994.

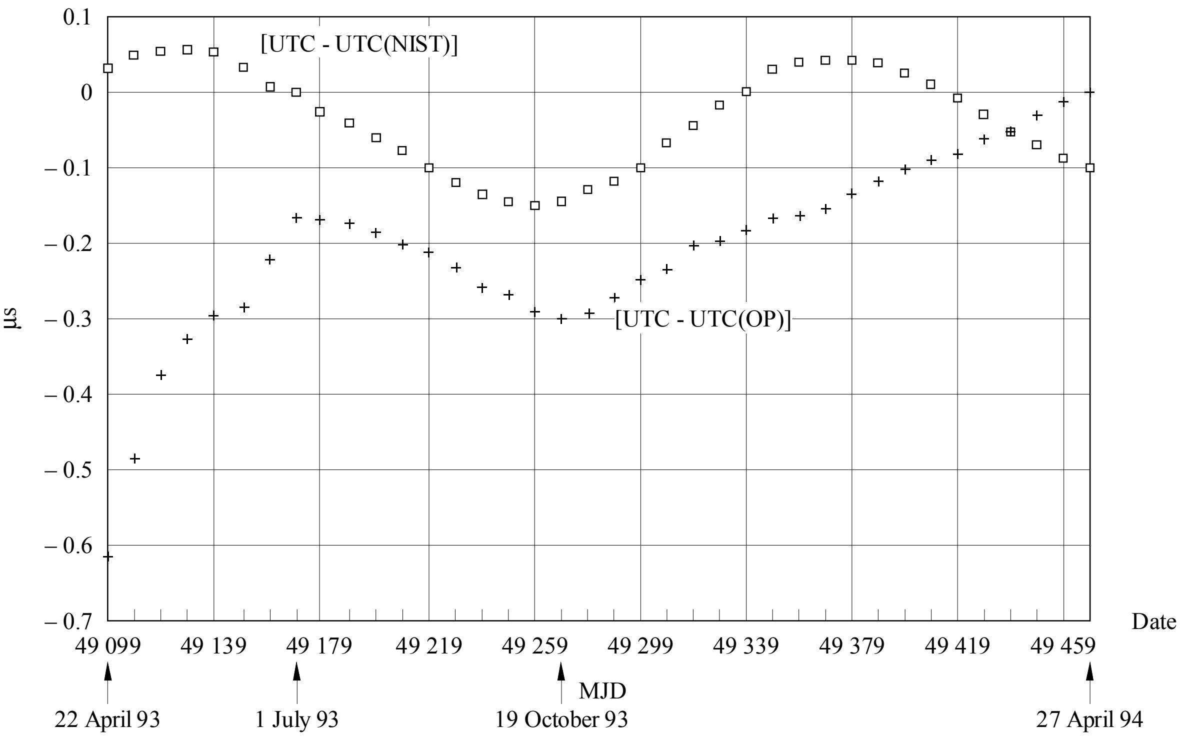


FIGURE 6.5

Timing data obtained from the comparison between UTC and UTC(OP) and UTC and UTC(NIST)

The UTC(OP), in Paris, is a hardware UTC derived from one single physical clock steered through a micro-phase-stepper. The change of the master clock in 1993 from an old-design HP unit to a HP 5071A unit is readily observed: stability is immediately improved. A frequency steering command was given in 1993 to bring UTC(OP) closer to UTC.

The UTC(NIST) kept at the NIST is a software UTC derived from an ensemble of physical clocks, and steered to UTC by software. This local UTC has several physical representations, obtained from hardware clocks, each steered every 6 minutes, through a microphase stepper. The UTC(NIST) makes slow and regular oscillations around UTC.

6.4 Dissemination of time scales

As already described above, a time scale can be derived only from a knowledge of the time difference between this time scale and another one, or from a physical clock, at a given date. Access to time scales is thus effected by the publication of time differences. The uncertainty of these values is generally better than 10 ns (1 σ).

Before considering particular examples, it is useful to note that an ensemble time scale can be disseminated by making a comparison with any other operating clock, even if this clock does not participate in the generation of the time scale, for it is sufficient to have a time link. It is thus important to distinguish between the generation and the dissemination of a time scale. In the extreme case TAI could be chosen to be the mean of a few ultra-stable clocks kept in a small number of laboratories, but the work carried out for its dissemination, *i.e.* the establishment of an international GPS network, would be exactly the same.

The dissemination of most time scales is performed by the publication of official documents, usually on paper sheets, but also via electronic mail.

Figure 6.6 reproduces the first page of one issue of the monthly *Bulletin H* produced by the LPTF, the French primary laboratory for time and frequency . It contains several tables, in particular one which gives the time comparison values between UTC(OP) and GPS time, and between UTC(OP) and three European Loran-C chains. It covers a one-month period.

OBSERVATOIRE DE PARIS — Bulletin H 317

LABORATOIRE PRIMAIRE DU TEMPS ET DES FREQUENCES

Laboratoire primaire désigné par le Bureau National de Métrologie

TABLEAU 1 - MESURES DE TEMPS RAPPORTEES A UTC(OP)

Date 1994 Mai	Date MJD	MESURES DE PHASE DES CHAINES DE LORAN-C UTC(OP) - SIGNAL à 10h UT: SYLT 7970-H µs	ESTARTIT 7990-Z µs	LESSAY 8940-H µs	MESURES DU TEMPS GPS à 14h UT: UTC(OP) -GPS -9 s + µs
1	49473	3.35	1.73	-0.27	0.070
2	49474	3.38	1.76	-0.28	0.074
3	49475	3.46	1.82	-0.29	0.072
4	49476	3.42	1.72	-0.25	0.077
5	49477	3.44	1.72	-0.27	0.062
6	49478	3.52	1.78	-0.21	0.084
7	49479	3.51	1.83	-0.23	0.088
8	49480	3.47	1.78	-0.23	0.092
9	49481	3.48	1.74	-0.21	0.114
10	49482	3.49 (1)	1.74	-0.16	0.111
11	49483	3.34	1.81	-0.17	0.090
12	49484	3.33	1.81	-0.17	0.087
13	49485	3.33	1.77	-0.17	0.087
14	49486	3.38	1.93	-0.17	0.099
15	49487	3.34	1.87	-0.11	0.096
16	49488	3.41	1.95	-0.18	0.099
17	49489	3.33	1.87	-0.11	0.074
18	49490	3.33	1.79	-0.11	0.087
19	49491	3.35	1.82	-0.10	0.072
20	49492	3.35	1.92	-0.06	0.082
21	49493	3.40	1.88	-0.07	0.084
22	49494	3.43	1.91	-0.11	0.078
23	49495	3.44	1.88	-0.10	0.072
24	49496	3.51	1.84	-0.09	0.061
25	49497	3.46	1.81	-0.05	0.063
26	49498	3.52	1.82	-0.07	0.067
27	49499	3.50	1.81	-0.08	0.073
28	49500	3.52	1.79	-0.06	0.074
29	49501	3.50	1.76	-0.03	0.087
30	49502	3.52	1.77	-0.05	0.070
31	49503	3.52	1.80	0.01	0.075

L P T F , Observatoire de PARIS, 61 Avenue de l'Observatoire, 75014 PARIS, FRANCE
Téléphone : (33-1) 40-51-22-21 - Télex : 270 776 - Téléfax : (33-1) 43-25-55-42
Internet: lptfop@opdaf1.obspm.fr - Span : opdaf1::lptfop

BIPM
- 9. JUIN 1994
Répondu le :

FIGURE 6.6

First page of *Bulletin H* (issue No 317), produced on a monthly basis at the LPTF, Paris, France

Figure 6.7 reproduces the two first pages of one issue of the weekly *IERS Bulletin-A*. It contains tables of values of comparison between UT1 and UTC and information about the polar motion.

IERS Bulletin-A

International Earth Rotation Service

NEOS Earth Orientation Bulletin

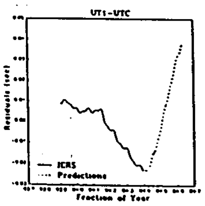
26 May 1994 — Incorporating U.S.N.O. Series 7 — Vol. VII No. 21

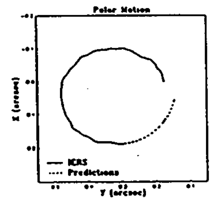
CONTENTS OF THIS ISSUE: PAGE

GENERAL INFORMATION:
MJD = Julian Date - 2 400 000.5 days
UT2-UT1 = 0.022 sin(2*pi*T) - 0.012 cos(2*pi*T) - 0.006 sin(4*pi*T) + 0.007 cos(4*pi*T)
where pi = 3.14159265... and T is the date in Besselian years.
TDT = TAI + 32.184 seconds = A.1 + 32.153 seconds
DUT1= (UT1-UTC) transmitted with time signals
= -0.2 seconds beginning 9 June 1994 at 0000 UTC
Beginning 1 July 1993:
TAI-UTC(BIPM) = 28.000 000 seconds
A.1(USNO)-UTC(USNO) = 28.034 381 7 seconds

A positive leap second will be introduced
in UTC on 30 June 1994.

U.S.N.O. DATA AND PREDICTIONS OF UT1-UTC AND POLAR MOTION
T(PLOT) = UT1-UTC - .2811 + (MJD-49315) * .00226





OBSERVATIONS:

NEOS VLBI Intensives

MJD	UT1-UTC (s)	s.e. (s)
49483.7962	-.123264	.000035
49485.7863	-.127547	.000030
49486.7837	-.129754	.000030

P designates preliminary values based on preliminary pole positions

NEOS VLBI

MJD	x (″)	s.e. (″)	y (″)	s.e. (″)	UT1-UTC (s)	s.e. (s)
49490.253	.18678	.00010	.29495	.00007	-.137987	.000004

VLBI from Jet Propulsion Laboratory

MJD	UT0-UTC (sec)	s.d. (sec)	Baseline
49490.010	-.14455	.00025	1463
49493.850	-.14673	.00005	1463

DELFT UNIVERSITY OF TECHNOLOGY SLR

MJD	x (″)	m.e (″)	y (″)	m.e. (″)
49488.00	.18917	.00015	.30282	.00013
49491.00	.18758	.00018	.29647	.00014

GLOBAL POSITIONING SYSTEM COMBINED SOLUTION

MJD	x (″)	m.e (″)	y (″)	m.e. (″)
49489.0	.18837	.00040	.29903	.00038
49490.0	.18802	.00035	.29698	.00041
49491.0	.18771	.00033	.29494	.00037
49492.0	.18741	.00023	.29289	.00036
49493.0	.18710	.00024	.29086	.00036
49494.0	.18677	.00027	.28882	.00038
49495.0	.18646	.00029	.28678	.00038

COMBINED EARTH ORIENTATION PARAMETERS:

IERS Rapid Service

	MJD	x (″)	error (″)	y (″)	error (″)	UT1-UTC (s)	error (s)
1994 May 17	49489	.18847	.00030	.29906	.00016	-.134786	.000076
18	49490	.18816	.00032	.29700	.00018	-.137299	.000157
19	49491	.18783	.00031	.29495	.00018	-.139949	.000165
20	49492	.18748	.00031	.29290	.00017	-.142710	.000165
21	49493	.18712	.00026	.29086	.00017	-.145522	.000167
22	49494	.18677	.00024	.28882	.00018	-.148303	.000363
23	49495	.18642	.00024	.28678	.00011	-.150971	.000440

PREDICTIONS:
The following formulas will not reproduce the predictions given below, but may be used to extend the predictions beyond the end of this table.

x = .0270 - .0279 cos A + .0756 sin A + .1938 cos C - .1141 sin C
y = .3224 + .0721 cos A + .0337 sin A - .1141 cos C - .1938 sin C
UT1-UTC = -.1024 - .00231 (MJD - 49495) - (UT2-UT1)

2

FIGURE 6.7

First two pages of *IERS Bulletin-A* (issue of 26 May 1994), produced on a weekly basis at the IERS, Paris, France

Figure 6.8 reproduces the three first sections of one issue of the monthly *Circular T* produced by the BIPM. It contains tables of the values of comparisons between UTC and UTC(k) for the 45 local representations of UTC, and between TAI and TA(k) for the 17 independent atomic time scales computed throughout the world. The BIPM also gives a daily estimation of comparisons between UTC and GPS time, and UTC and GLONASS time.

BIPM ISSN 1143-1393

BUREAU INTERNATIONAL DES POIDS ET MESURES

Circular T 76 (1994 May 25)

1 - Coordinated Universal Time UTC. Computed values of UTC-UTC(k) (1).

(From 1993 July 1, 0h UTC, to 1994 July 1, 0h UTC, TAI-UTC = 28 s)
(From 1994 July 1, 0h UTC, until further notice, TAI-UTC = 29 s)

Date 1994 0h UTC		Mar 28	Apr 7	Apr 17	Apr 27
MJD		49439	49449	49459	49469
Laboratory k		UTC-UTC(k) (Unit = 1 microsecond)			
AOS (Borowiec)		-1.153	-1.377	-1.520	-1.701
APL (Laurel)		1.249	1.181	1.126	1.061
AUS (Canberra)		0.498	0.476	0.447	0.408
BEV (Wien)		-	-	-	-
CAO (Cagliari)		-6.872	-7.066	-7.233	-7.483
CH (Bern)		1.929	1.921	1.826	1.693
CRL (Tokyo)		2.053	2.024	2.010	2.025
CSAO (Lintong)		-0.477	-0.452	-0.407	-0.402
CSIR (Pretoria)		-3.061	-2.924	-2.865	-2.826
FTZ (Darmstadt)		0.048	0.052	0.161	0.220
IEN (Torino)		0.065	0.103	0.131	0.177
IFAG (Wettzell)		-0.598	-0.575	-0.569	-0.522
IGMA (Buenos Aires)		-3.14	-3.14	-3.13	-3.15
INPL (Jerusalem)		-1.135	-1.280	-1.398	-1.474
JATC (Lintong)		-1.091	-1.283	-0.397	0.365
~~KRIS~~ (Taejon)		-0.276	-0.289	-0.264	-0.251
LDS (Leeds)		-0.282	-0.281	-0.323	-0.356
MSL (Lower Hutt)		-0.434	-0.431	-0.388	-0.346
NAOM (Mizusawa)		-1.436	-1.477	-1.513	-1.539
NAOT (Tokyo)		-0.661	-0.876	-1.035	-1.262
NIM (Beijing)		7.78	7.81	7.78	7.80
NIST (Boulder)		-0.051	-0.068	-0.086	-0.094
NMC (Sofiya)		-	-	-	-
NPL (Teddington)		0.119	0.116	0.114	0.113
NPLI (New-Delhi)	(2)	-3.12	-3.22	-	-3.18
NRC (Ottawa)		5.265	5.367	5.468	5.567
NRLM (Tsukuba)		-9.641	-9.937	-10.233	-10.521
OMH (Budapest)		6.489	6.510	6.502	6.559
ONBA (Buenos Aires)		5.65	5.57	5.70	5.48
ONRJ (Rio de Janeiro)		-13.877	-	-	-
OP (Paris)		-0.047	-0.029	-0.010	0.005
ORB (Bruxelles)		-1.673	-1.712	-1.666	-1.755
PKNM (Warszawa)		0.454	0.373	0.241	0.212
PTB (Braunschweig)		2.748	2.753	2.754	2.772
RC (Habana)	(3)	-2.36	-3.00	-3.08	-2.80
ROA (San Fernando)		2.615	2.610	2.632	2.637
SCL (Hong Kong)		0.034	0.107	0.177	0.424
SMT (Stockholm)		0.065	0.085	0.086	0.067
SO (Shanghai)		2.14	-	2.16	1.78
SU (Moskva)		-3.375	-3.461	-3.548	-3.624
TL (Chung-Li)		-3.106	-3.049	-2.985	-2.914
TP (Praha)		-1.147	-1.135	-1.098	-1.069
TUG (Graz)		4.481	4.564	4.643	4.739
USNO (Washington DC)(USNO MC)		0.045	0.051	0.053	0.057
VSL (Delft)		0.094	0.132	0.166	0.174

PAVILLON DE BRETEUIL F - 92312 SEVRES CEDEX
TEL CENTRAL + 33 1 45 07 70 70 TELEX BIPM 631351 F TELECOPIE + 33 1 45 34 20 21

BIPM . T 76 (2)

2 - International Atomic Time TAI and local atomic time scales TA(k).

The following table gives the computed values of TAI-TA(k) (1).

Date 1994 0h UTC		Mar 28	Apr 7	Apr 17	Apr 27
MJD		49439	49449	49459	49469
Laboratory k		TAI-TA(k) (Unit = 1 microsecond)			
APL (Laurel)		2.712	2.644	2.589	2.524
AUS (Canberra)		-50.849	-51.020	-51.106	-51.273
CH (Bern)		-75.231	-75.059	-74.894	-74.767
CRL (Tokyo)		36.656	37.065	37.496	37.948
CSAO (Lintong)		14.992	14.887	14.803	14.678
F (Paris)		127.851	128.227	128.604	128.987
INPL (Jerusalem)		-	-196.410	-198.459	-200.492
JATC (Lintong)		9.415	10.044	10.708	11.408
KRIS (Taejon)		-3.486	-3.279	-3.054	-2.811
NIM (Beijing)		-8.73	-8.68	-8.70	-8.66
NISA (Boulder)	(4)	-45111.230	-45111.631	-45112.029	-45112.417
NRC (Ottawa)		21.334	21.436	21.537	21.636
PTB (Braunschweig)		-360.652	-360.647	-360.646	-360.628
RC (Habana)	(3)(5)	-325.85	-326.53	-326.66	-326.42
SO (Shanghai)		-45.43	-	-45.40	-45.81
SU (Moskva)	(6)	27246.625	27246.539	27246.452	27246.376
USNO (Washington DC)	(7)	-34695.858	-34696.529	-34697.211	-34697.888

3 - Notes on sections 1 and 2.

(1) Values UTC-UTC(k) and TAI-TA(k) are published within 1 ns except for laboratories which are not linked through GPS common views.

(2) NPLI. MJD UTC-UTC(NPLI)

MJD	UTC-UTC(NPLI)
49419	-3.29
49429	-3.03

(3) RC . MJD UTC-UTC(RC) TAI-TA(RC) - 18 s

MJD	UTC-UTC(RC)	TAI-TA(RC) - 18 s
49419	-2.78	-326.18
49429	-2.54	-325.99

(4) NIST. TA(NISA) designates the scale AT1 of NIST.

(5) RC . Listed values are TAI-TA(RC) - 18 seconds.

(6) SU . Listed values are TAI-TA(SU) - 2.80 seconds.

(7) USNO. TA(USNO) designates the scale A1(MEAN) of USNO.

FIGURE 6.8

First two pages of *Circular T* (issue of 25 May 1994), produced on a monthly basis at the BIPM, Sèvres, France

For other time scales, such as GPS time and GLONASS time, dissemination is realised in real time through observations of the satellites which transmit it. It may be necessary to filter the measurements to remove observational noise and intentional degradation. Deferred-time access is obtained by specific publications produced by the USNO [Series 4] (see Figure 6.9), the BIPM [*Circular T*], and also by the NIST time and frequency services.

DAILY TIME DIFFERENCES, SERIES 4, NO. 1426 (CONTINUED)

GLOBAL POSITIONING SYSTEM (GPS)
BLOCK I AND BLOCK II SATELLITES

VALUES PRESENTED BELOW FOR NAVSTAR GPS SATELLITES ARE THE RESULT OF A LINEAR FIT THROUGH APPROXIMATELY 130 DATA POINTS REFERRED TO THE BEGINNING OF THE TRACKING PERIOD. TRACKING PERIODS START ON THE MINUTE AND RANGE FROM TWO TO THIRTEEN MINUTES.

GPS TIME IS AHEAD OF UTC BY NINE SECONDS. UNIT - ONE NANOSECOND

		NAVSTAR 10 PRN12		NAVSTAR 13 PRN02		NAVSTAR 14 PRN14		NAVSTAR 15 PRN15		NAVSTAR 16 PRN16	
	MJD	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME
MAY 22	49494	0	(2102)	-96	(2005)	-40	(2135)	-7	(1901)	41	(0254)
23	49495	-13	(2058)	-17	(2001)	-87	(2135)	-16	(1857)	-53	(0251)
24	49496	-6	(2054)	-52	(1957)	67	(2123)	-9	(1853)	-34	(0247)
25	49497	-5	(2050)	-8	(1953)	-49	(2119)	-15	(1849)	61	(0243)
26	49498	2	(2046)	-137	(1949)	-29	(2115)	-20	(1845)	17	(0239)
27	49499	-1	(2042)	-34	(1945)	-27	(2111)	-8	(1841)	-71	(0235)
28	49500	9	(2038)	-7	(1941)	-42	(2107)	-5	(1837)	27	(0231)
29	49501	-3	(2034)	38	(1937)	-79	(2103)	-5	(1833)	88	(0227)
30	49502	0	(2030)	249	(1933)	-12	(2059)	-2	(1829)	10	(0223)
31	49503	8	(2026)	0	(1930)	148	(2055)	4	(1825)	-109	(0219)

		NAVSTAR 17 PRN17		NAVSTAR 18 PRN18		NAVSTAR 19 PRN19		NAVSTAR 20 PRN20		NAVSTAR 21 PRN21	
	MJD	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME	MC-GPS	GPS TIME
MAY 22	49494	73	(0607)	-9	(2327)	67	(1717)	18	(0954)	37	(0818)
23	49495	-6	(0603)	-11	(2323)	29	(1713)	-10	(2351)	44	(0814)
24	49496	-28	(0559)	119	(2319)	81	(1709)	-17	(2347)	4	(0810)
25	49497	-2	(0555)	141	(2315)	97	(1705)	22	(2356)	6	(0806)
26	49498	-51	(0551)	1	(2311)	-55	(1701)	-5	(2339)	176	(0802)
27	49499	20	(0547)	216	(2307)	70	(1657)	0	(2335)	-85	(0758)
28	49500	12	(0543)	-5	(2303)	11	(1653)	16	(2331)	32	(0754)
29	49501	26	(0540)	-71	(2259)	-61	(1649)	13	(2327)	69	(0750)
30	49502	-59	(0540)	190	(2358)	-25	(1645)	14	(2323)	91	(0746)
31	49503	138	(0541)	-30	(2354)	101	(1641)	29	(2332)	62	(0742)

FIGURE 6.9

Second page of *USNO Series 4* (issue No 1426), produced on a weekly basis at the USNO, Washington, D.C., USA

6.5 Conclusions

In 1991, the International Astronomical Union clearly specified in terms of general relativity the framework within which time scales should be defined. A realisation of Terrestrial Time, as explicitly mentioned in the IAU Resolution, is International Atomic Time, TAI, which is obtained from a combination of the readings of atomic clocks kept on the Earth.

While TAI is the international reference for timing, many other time scales are regularly computed and used for scientific purposes. Besides keeping local representations of UTC, the laboratories which compute these scales have to design algorithms for the generation of free-running and independent time scales based on data collected on site. Development of algorithms inevitably leads to a need to write an equation of definition, in the form of a weighted average, and to settle procedures for the determination of clock weight and the prediction of clock frequency. Many sophistications are possible, but the actual choices are guided by the purposes the time scale is intended to serve and by the noise affecting timing data.

In 1993, world-wide most stable time scales reached stabilities better than 1×10^{-14} for averaging times of the order of several weeks. Achieved accuracies are limited by the accuracy of the best primary frequency standards, and are, at present, characterised by an uncertainty (1 σ) of order 1×10^{-14}. Improvements in performance are rapid: it is probable that accuracies of order some parts in 10^{16}, for realisation of the SI second, and several hundreds of picoseconds, for time comparisons, will be available in the year 2000.

Though the second is defined in atomic terms and time scales are generated from atomic clocks, time retains its close relation with astronomy: the international reference time scale is the purely atomic TAI, but coherence with the Earth rotation has been maintained by the production of UTC. The 21th century may see the relation with astronomy again reinforced through the use of millisecond pulsars for monitoring the long-term stability of TAI [Petit *et al.*, 1992].

Note – Laboratory acronyms and locations can be found in Table 3, page 20 and 21 of the Annual Report of the BIPM Time Section, Volume 6, available on request from the BIPM, Pavillon de Breteuil, 92312 Sèvres Cedex, France.

References

ALLAN D.W. [1987] Time and frequency (time domain) characterization, estimation, and prediction of precision clocks and oscillators. *IEEE Trans. Ultr. Ferr. Freq. Control*, UFFC-34, 647-654.

ALLAN D.W. and ASHBY N. [1986] *Relativity in Celestial Mechanics and Astrometry*. Ed. Kovalevsky J. and Brumberg V.A., Reidel Dordrecht, 299-313.

ALLAN D.W. and DAAMS H. [1975] Picosecond Time Difference Measurement System. *Proc. 29th FCS*, 404-411.

ALLAN D.W. and WEISS M.A. [1980] Accurate Time and Frequency Transfer during Common-View of a GPS Satellite. *Proc. 34th FCS*, 334-346.

ALLAN D.W. *et al.*, [1974] The National Bureau of Standards Atomic Time Scale: Generation, Stability, Accuracy, and Accessibility. *NBS Monograph 140*, ed. Byron Blair, Chapter 9.

AZOUBIB J., GRANVEAUD M. and GUINOT B. [1977] Estimation of the scale unit duration of time scales. *Metrologia*, Vol. 13, 87-93.

BARNES J.A. [1982] Time Scales Algorithms Using Kalman Filters – Insights from Simulation. *Proc. 2nd Symposium on Atomic Time Scales Algorithms*, Sect. 15, 42.

BAUMONT F., GRUDLER P., VEILLET C., WIANT J., LEWANDOWSKI W. and PETIT G. [1993] A preliminary report on the comparison of LASSO and GPS time transfer. *Proc 7th EFTF*, 641-646.

BIPM [1991] Le Système International d'Unités, SI, 6e Edition. Bureau International des Poids et Mesures.

BIPM [1993] Annual Report of the BIPM Time Section, Vol. 6, 129 pages.

BREAKIRON L.A. [1991] Timescale algorithms combining cesium clocks and hydrogen masers. *Proc. 23th PTTI Meeting*, 297-305.

BRUMBERG V.A. [1991] *Essential Relativistic Celestial Mechanics.* Adam Hilger, Bristol.

CCDS Report [1980] BIPM Com. Cons. Déf. Seconde, Vol. 9, p S 15.

CCDS Report [1993] BIPM Com. Cons. Déf. Seconde, Vol. 12, p S 69.

DALY P., KOSHELYAEVSKY N.B., LEWANDOWSKI W., PETIT G. and THOMAS C. [1992] Comparison of GLONASS and GPS Time Transfers. *Proc 6th EFTF*, 249-252.

DE JONG G. [1993] Two-Way Satellite Time Transfer: Overview and Recent Developments. *Proc. 25th PTTI Meeting*, 101-117.

DE MARCHI A. [1988] The Accuracy of Commercial Cesium Beam Frequency Standards. *Frequency Standards and Metrology,* ed. A. De Marchi, Springer-Verlag, Berlin, 52-56.

FEESS W.A., HOLTZ H., SATIN A.L. and YINGER C.H. [1991] Evaluation of GPS/UTC steering performance. *Proc. 23th PTTI Meeting*, 35-46.

GRANVEAUD M. [1986] Echelles de temps atomique, Monographie du Bureau National de Métrologie. Ed. Chiron.

GUINOT B. [1986] Some properties of algorithms for atomic time scales. *Metrologia*, 1987, Vol. 24, 195-198.

GUINOT B. [1988] Atomic time scales for pulsar studies and other demanding applications. *Astronomy and Astrophysics*, Vol. 192, 370-373.

GUINOT B. and THOMAS C. [1988] Establishment of International Atomic Time. *Annual Report of the BIPM Time Section*, Vol. 1, D1-D22.

IAU [1991] IAU transactions Vol. XXIB. *Proc. 21st Gen. Assembly Buenos Aires*, Kluwer Acad. Publ., Dordrecht, Boston, London.

IAU [1992] Information Bulletin 67, p. 7.

JONES R.H. and TRYON P.V. [1983] Estimating time from atomic clocks. *Journal of Research of the NBS*, Vol. 88, 1, 17-24.

JONES R.H. and TRYON P.V. [1987]Continuous Time Series Models for unequally Spaces Data Applied to Modeling Atomic Clocks. *SIAM J. Sci. Stat. Comput.*, Vol. 8, 71-81.

KIRCHNER D., THYR U., RESSLER H., ROBNIK R., GRUDLER P., BAUMONT F., VEILLET C., LEWANDOWSKI W., HANSON W., CLEMENTS A., JESPERSEN J., HOWE D., LOMBARDI M., KLEPCZYNSKI W., WHEELER P., POWELL W., DAVIS A., UHRICH P., TOURDE R. and GRANVEAUD M. [1991] Comparison of Two-Way Satellite Time Transfer and GPS Common-View Time Transfer Between OCA and TUG. *Proc. 23rd PTTI*, 71-88.

LEWANDOWSKI W. and THOMAS C. [1991] GPS Time Transfer. *Proc. IEEE Special Issue on Time*, Vol. 79, 991-1000.

LEWANDOWSKI W., MOUSSAY P., CHERENKOV G.T., KOSHELYAEVSKY N.B. and PUSHKIN S.B. [1993] GLONASS common-view time transfer, *Proc. 7th EFTF*, 147-151.

LUCK J.M. [1979] Comparison and coordination of time scales. *Proc. Astronomical Society of Australia*, Vol. 3, 5-6, 357-363.

MISNER C.W., THORNE K.S. and WHEELER J.A. [1973] *Gravitation.* W.H. Freeman and Company, San Francisco.

PETIT G., TAVELLA P. and THOMAS C. [1992] How can Millisecond Pulsars improve the Long-Term – Stability of Atomic Time Scales?. *Proc. 6th EFTF*, 57-60.

PETIT G. and WOLF P. [1994] Relativistic theory for picosecond time transfer in the vicinity of the Earth. *Astronomy and Astrophysics*, Vol. 286, 971-977.

PERCIVAL D.B. [1978] The U.S. Naval Observatory Time Scales. *IEEE Trans. Instrum. Meas.*, IM-27, 376-385.

STEIN S.R. [1988] Kalman Ensembling Algorithm: Aiding Sources Approach. *Proc. 3rd International Time Scale Algorithm Symposium*, 345-358.

STEIN S.R. [1992] Advances in time-scale algorithms. *Proc. 24th PTTI Meeting*, pp. 289-302.

STEIN S.R., GIFFORD G.A. and BREAKIRON L.A. [1989] Report on the Timescale Algorithm Test Bed at USNO. *Proc. 21st PTTI Meeting*, 269-288.

STEIN S.R., GLAZE D., LEVINE J., GRAY J., HILLIARD D. and HOWE D. [1982] Performance of an Automated High Accuracy Phase Measurement System. *Proc. 38th FCS*.

TAVELLA P., AZOUBIB J. and THOMAS C. [1991] Study of the Clock-Ensemble Correlation in ALGOS Using Real Data. *Proc. 5th EFTF*, 435-441.

TAVELLA P. and PREMOLI A. [1994] Estimating the instabilities of *N* clocks by measuring differences of their readings. *Metrologia*, Vol. 30, 479-486.

TAVELLA P. and THOMAS C. [1990a] Time Scale Algorithm: Definition of Ensemble Time and Possible Uses of the Kalman Filter. *Proc. 22nd PTTI Meeting*, 157-170.

TAVELLA P. and THOMAS C. [1990b] Study of the correlations among the frequency changes of the contributing clocks to TAI. *Proc. 4th EFTF*, 527-541.

TAVELLA P. and THOMAS C. [1991a] Comparative Study of Time Scale Algorithms. *Metrologia*, Vol. 28, 57-63.

TAVELLA P. and THOMAS C. [1991b] Report on correlations in frequency changes among the clocks contributing to TAI. Rapport BIPM-91/4, 50 pages.

VARNUM F.B., BROWN D.R., ALLAN D.W., and PEPPLER T.K. [1987] Comparison of time scales generated with the NBS ensembling algorithm. *Proc. 19th PTTI Meeting*, 13-24.

WEI GUO [1992] A study of atomic time scale stability. *Proc. 46th FCS*, 151-156.

WEISS M.A., ALLAN D.W. and PEPPLER T.K. [1989] A Study of the NBS Time Scale Algorithm. *IEEE Trans. Instr. Meas.*, IM-38, 631-635.

WEISS M.A. and WEISSERT T. [1991] AT2, A new time scale algorithm: AT1 plus frequency variance. *Metrologia*, Vol. 28, 65-74.

WEISS M.A. and WEISSERT T. [1994] Sifting through Nine Years of NIST Clock Data with TA2. *Metrologia*, Vol. 31, 9-19.

YOSHIMURA K. [1980] Calculation of unbiased clock-variances in uncalibrated atomic time scale algorithms. *Metrologia*, Vol. 16, 133-139.

Chapter 7

Uses of frequency sources

Contents

7.1 Use of frequency sources in science and technology

In what follows, the acronym FS means a Frequency source, or a Frequency Standard, i.e. a device, as described in Chapters 1 and 2 (Part A), which is able to provide an electrical signal whose output frequency can be assumed to satisfy all the requirements needed by the application of interest. Such requirements can be accuracy, stability, insensitivity to environment, spectral purity, etc., as discussed in Chapters 4 and 5. No assumption is made, for the moment, as to the type of the FS, since the choice is dictated by the application and other requirements.

Of all the man made devices, FS's are unique in the sense that FS's offer usually the best compromise between accuracy, reliability, cost, etc. This statement requires some explanations and examples. At a given level of accuracy, for example 10^{-7}, a FS or a frequency measuring instrument costs less than 1/100 of the equivalent accuracy devices for length or mass. In technical applications, the useful lives of commercial devices for FS's range from 5 to 20 years, depending on the type. Power requirements too can be very small; the FS's used in crystal wrist watches or similar devices run with a power less than one microwatt. Telecommunications of every type rely heavily on FS's; a cellular mobile phone is fitted with 4 to 5 FS's, while every TV-set or computer has inside at least one FS.

In measurement technology it is usually convenient to convert, via a transducer, the quantity of interest, whatever its nature – voltage, pressure, humidity, speed, etc. – into a frequency or into a time interval that is finally measured using a FS. This procedure offers relevant gains concerning accuracy, cost and ease of use, because the greatest accuracy and precision can be obtained from a FS at the smallest cost. Finally, frequency standards are unique in solving a host of problems, such as the measurement of mass, velocity, or acceleration or position of a remote spacecraft.

7.2 Metrology

The aim of this section is to describe the present and future foreseeable relations between time and frequency metrology, others metrologies and the fundamental constants. The International System of Units, designed as SI (Système International d'Unité) is based on seven basic quantities, each one having a standard unit, described by a definition. Linked to these basic units is the countless set of derived units (velocity, resistance, heat transfer rate, specific weight, etc.) used in science and technology. Concerning the quantity time, the unit -the second - and its definition were introduced in Chapters 1 and 2.

7.2.1 Accuracy comparison between the standard of time and those of the other basic quantities

As seen in Chapters 1 and 2, the time standard is derived directly, with a minimum of assumptions, from a fundamental constant and from some properties of matter. As again seen in Chapter 1 and 2, there are indeed sources of errors, but with suitable techniques the second can be reproduced in each metrological laboratory with an uncertainty between 1×10^{-13} and 1×10^{-14}. For a number of reasons, not covered here, the accuracy available for the second greatly exceeds those achievable in the experimental realisations of the other units. Table 7.1 and Figure 7.1 depict the current situation.

TABLE 7.1

Relative uncertainties in the SI units realisation

Base Unit	m	kg	s	A	K	mol	cd
Relative uncertainty	10^{-11}	10^{-8}	10^{-14}	10^{-7}	10^{-6}	10^{-6}	10^{-3}

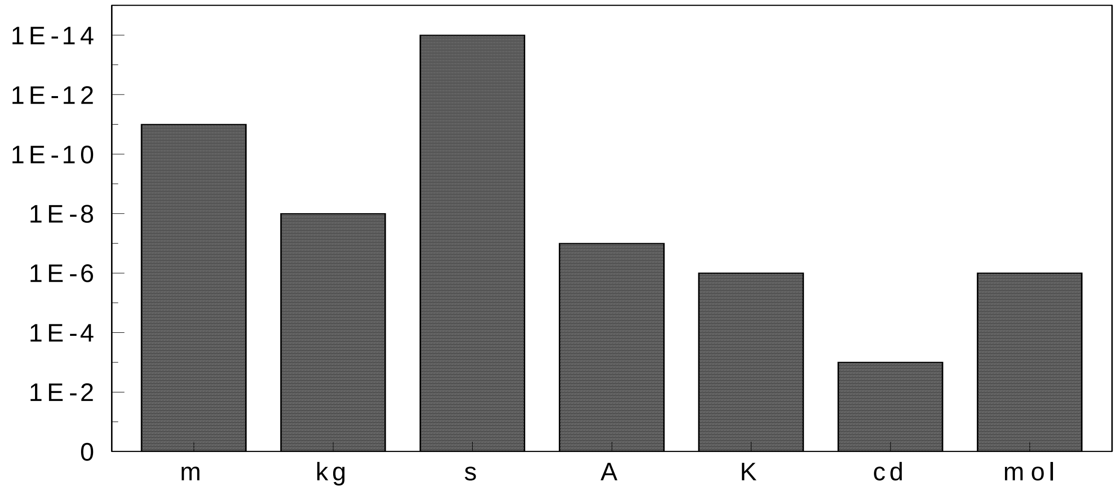


FIGURE 7.1

Relative uncertainty in the SI units realisation

The fact that a FS can be, at a given level of accuracy, far simpler and less expensive than other standards, partly explains what was indicated in the previous section concerning the interest to first transform any quantity into a frequency or into a time interval, to be then measured with the techniques seen *e.g.,* in Chapters 3 and 4 or outlined in Section 7.6.3.

7.2.2 Relations between the unit of time and other units

In the past, the Metric System was based also on a number of artefacts such as the bar giving the meter, cell for the volt or a metallic resistor for the ohm, and so on. Currently, one unit – that of mass – is still realised via an artefact, and some other units, fundamental or derived, are in progress towards a definition linked to Nature (Fundamental Constants, Laws of Physics) securing the stability and permanence of the definition. Indeed this effort has been already successful for many of the base standards.

An important step was, in 1983, the change of the definition of the metre, derived now directly from the second by a measurement and a definition of the value of the velocity of light.

The definition of the metre

In 1983, the Conférence Internationale des Poids et Mésures, adopted the following definition for the metre: *"the metre is the length of the path travelled by light in a vacuum during a time interval of 1/299 792 458 of a second."*

The unit of length is consequently now linked to the definition of the second and since for c an "exact" value was adopted, the accuracy available for the unit of time ($10^{-13} \rightarrow 10^{-14}$) can now, in principle, be transferred to the unit of length.

The maintenance of the electrical base unit, the ampere, relies on the representations of the volt and of the ohm, given by the value of two fundamental constants – the Josephson constant $K_j = 2e/h$ and the von Klitzing constant $R_k = h/e^2$ – plus, for the unit of voltage, a measurement of the frequency of a microwave signal. As a consequence, this electrical quantity is directly linked to a standard FS. The experiments used for the volt and for the ohm are the Josephson effect and the Quantum Hall effect respectively.

Theories for establishing a quantum standard of current, being directly related to the frequency and to the electron charge have also been presented. Details on these new standards can be found in [Pöpel, 1992; Hartland,1992].

The Josephson effect

The Josephson effect links directly a frequency to a voltage, by the ratio between *e*, the electron charge, and *h*, the Planck constant. This cryogenic phenomenon, foreseen theoretically in 1962, has fundamentally changed the electrical metrology; a voltage can be measured with an accuracy of about 10^{-10}, using a frequency counter. Conversely, if a voltage is adequately known with respect to the fundamental SI units, the ratio e/h, can be derived with unprecedented accuracy.

The unit of luminous intensity, the candela, was changed from an artefact in 1979, and is now derived from an electrical power measurement, the frequency of a laser, and geometric measurements. Consequently also the candela is linked, albeit indirectly, to a FS.

Regarding mass, research is underway to link the mass to electrical quantities, space, velocity and time. This bold enterprise is based on the concept that the power, in a given system, must be the same whether calculated by mechanical quantities (force and velocity) or electrical quantities (voltage and current).

As a final outcome it could be that all the basic units, with the exception of the quantity of substance, the mole, and thermodynamic temperature, the Kelvin, could be linked directly to Time and Frequency Metrology and that a FS will become the basic ingredient for the realisation of a number of fundamental and derived units [Petley, 1988].

7.3 Fundamental and applied physics

Most of the research activities on fundamental and applied physics are now based on FS devices or techniques. The range of applications is wide and covers: Physical laws, applied physics, astronomy, validation of post Newtonian theories, validation and application of special and general relativity, geodesy, geopotential and geophysics. FS devices and techniques play a fundamental role too in geoscience from space, such as oceanography, climatology and, in general, in remote sensing. Only a few of these FS applications will be outlined in what follows.

7.3.1 g, Gravity acceleration

The value of **g** is obtained in the laboratory and in the field, by measuring a free fall of a body. The body is a square corner cube reflector falling in vacuum; its trajectory is tracked using laser interferometer, since the falling retroreflector forms the variable arm of a Michelson interferometer. The laser can be stabilised with reference to a FS, and the time tagging of the fall is performed using an atomic clock. The local value of **g** can be obtained with an accuracy of 10^{-9}.

7.3.2 GM, Gravitational constant times earth mass

In satellite orbit calculations, the quantity of interest is the product **GM** and not the individual values for the gravitational constant **G**, and the Earth's mass **M**. An accurate value of the product **GM** is obtained via the Kepler's third law, launching a satellite around the Earth and measuring the orbit parameters. Kepler's third law can be written as:

$$G(M_1 + M_2) = k \cdot a^3/P^2$$

where P is the orbit period, a is the orbit semi-axis and M_1 and M_2 are the Earth and satellite masses, respectively.

The mass M_2 can be obviously disregarded, the semi-axis a can be measured by time interval measurements (laser ranging), with an accuracy of about 10^{-9}, the period P, using Doppler or laser orbitography, atomic clocks and TAI, can be reckoned with an accuracy of about $10^{-8} \rightarrow 10^{-9}$. The value of the product **GM** can be obtained with an accuracy of about 10^{-9}. Of some interest is the fact that both **G** and **M** are separately known with errors of $10^{-4} \rightarrow 10^{-5}$.

7.3.3 The Earth gravitational field

The value and the distribution of the geopotential around the Earth reflects the mass distribution inside the Earth itself. These quantities are now known by observation of the anomalies in the orbits of some special satellites, inserted in well defined circular orbits. The instantaneous positions of these satellites are determined by measuring time intervals or frequency variations, i.e., by laser ranging or Doppler measurements.

The real shape of the geoid is thus available; a very powerful implementation in some geodetic satellites is the presence of an on-board radar altimeter, giving the small scale topography of continents and seas. By suitable treatment of the radar echoes, information about wind strength, wave heights, current locations, etc. can be deduced.

7.3.4 Very Long Baseline Interferometry (VLBI) and Quasi-VLBI

The scientific application requiring the utmost short term instability from a FS is the VLBI where, in two remote stations they measure the arrival time of pulses coming from a radio star. These measurements are subsequently compared and correlated. In this application the clocks to be used must have a frequency instability of the order of 10^{-14} for the duration of a measurement, typically ranging from 10 minutes to a few hours. The best FS for this task are the hydrogen masers; they are able to provide frequency stabilities of about 10^{-15} over one hour. VLBI data are instrumental in measuring continental drifts, the position and shape of extra-galactic radio sources, with an accuracy of 10 nanoradians, the wandering of the Pole, the small scale variations in the rotational speed of the Earth, and other astronomical and geophysical data.

A similar technique, called Quasi-VLBI, uses, instead of a natural radio source, such as a radio star, the coherent radio signals coming from an artificial satellite. This latter technique is useful in geodesy and surveying and in accurate space navigation.

7.4 Positioning and navigation

When one embarks an aeroplane, one should be aware that it is fitted with at least two dozen devices providing communication, localisation, navigation, hazard warning, etc., that are based on a clock or on a FS. Figure 7.2 depicts a typical situation, with a list of equipment. For each system, the accuracy needed on board or at satellites or at ground stations, is also given.

The requirements are on the average not very stringent, even if sometimes a temperature controlled or compensated quartz oscillator is specified, with the exception of navigation systems . In this very important feature, the quality needed for the clocks, in a very few cases for those on board, but always in the ground structure or in the satellites, must approach the performances of the FS that are used in the metrological laboratories.

Since this application of FS systems and methods is very crucial for travel efficiency and safety, it deserves special treatment, and since very few travellers are aware that the techniques involved rely on clocks. In what follows, the fundamental methods of radionavigation are outlined and the frequency and time measurement involved are listed and in some cases described.

The aim of this section is to provide elements and facts leading to understand why accurate FS are needed for navigation.

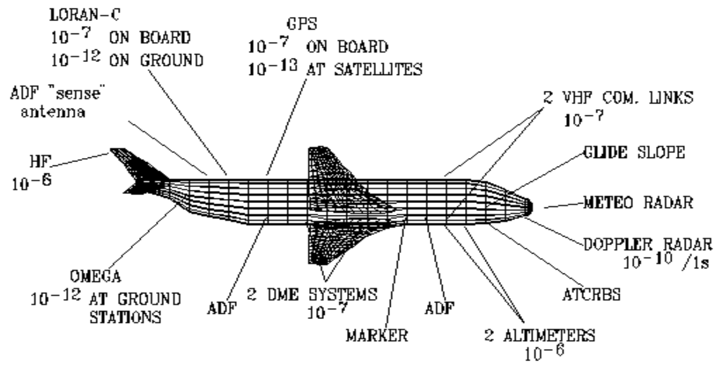


FIGURE 7.2

Aeroplane devices using time and frequency sources

Surveying positioning and navigation, among the oldest arts of Mankind, recently underwent a drastic change. For millennia, indeed, navigation was based on angle measurements and now, since the introduction of FS, most of the methods depends on distance or relative velocity measurements. Distances are obtained by time of flight of an electromagnetic pulse. Positions are obtained by intersections of lines or surfaces of position, whose shape and location, expressed in some adopted reference system, are given by measurements of:

- absolute time,
- time of flight,
- differences of time of arrival,
- phase differences,
- frequency variations.

All these measurements are performed using time and frequency equipment.

The geometric figures involved are lines, cones, circumferences, spheres, hyperbolae or hyperboloids. It is convenient to study the different methods and consequently the different uses of FS devices, using as a guideline the geometric figure considered, called place of position.

7.4.1 Conical navigation

With reference to Figure 7.3, all the points receiving a signal with the same Doppler effect,(see insert), must lie on the surface of a cone, having:

- its vertex V in the instantaneous position of the satellite,
- its axis tangent to the trajectory, at the vertex position,
- its semi-aperture angle alpha given by cos (alpha) = V_r/V_s, where V_r and V_s are, respectively, the relative velocity, measured at point P and the satellite velocity along its orbit.

The satellite transmits its position and all its orbit parameters, consequently the position of the vertex is known, V_s and its orientations are known, along the time at which the satellite was at the apex of the cone.

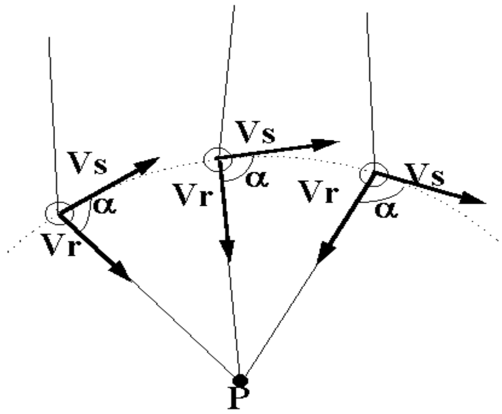


FIGURE 7.3

Principle of conical navigation

The value of Vr is measured at point at rest P, by measuring frequency with a frequency counter, driven by a suitable FS. It must be noted that all the points aligned with *V* and P are measuring the same Doppler effect and consequently the same applies for all the points over the cone surface.

The whole operation is repeated at later times, using the same satellite, resulting in a number of cones intersecting, in principle at point P. Thus, in the implementation of the conical method, two FS are needed, the first in the spacecraft, the second in the unknown point, plus a frequency measurement system to measure the Doppler effect in P.

The *Doppler effect* is the systematic variation of the frequency of any wave (acoustic, electromagnetic) propagating between two points in relative motion. With reference to Figure 7.3, let:

V_s	=	the velocity of a satellite along its orbit
F_s	=	the Frequency of an electromagnetic wave as emitted by the satellite
λ	=	the corresponding wavelength
V_1	=	the relative velocity between the satellite S and a point on the Earth, supposed at rest
F_r	=	the frequency received at P. As a first approximation, $F_r = F_s \pm V_a / \lambda$, where the sign + holds if V_r is positive (S is approaching P) and – holds if V_r is negative

Since the users are only receiving, their number is unbounded.

Attention is driven to the fact the FS devices used on the satellites and at ground stations are fully independent devices, i.e. they are not synthesised nor synchronised by other means, nor are they inside a network. Consequently all the measurements are absolute ones and the individual FS's must be independently linked or calibrated to a SI second source. Conical navigation is performed by some positioning satellite systems, such as ARGOS.

7.4.2 Circular or spherical navigation

Circular or spherical navigation follow basically the same principle, respectively in the plane and in the space. For simplicity, let us assume the simplest case of circular navigation in a plane.

With reference to Figure 7.4, let A and B be two points at rest or in motion, but with known positions. Point P measures its distances (the radii of circumferences) via a determination of the time of flight of an electromagnetic pulse, using the assumption that the signal propagates with the velocity of light. There are ambiguities, since two circumferences cross in two points, but it can be removed with other means or by a measurement against a third point C.

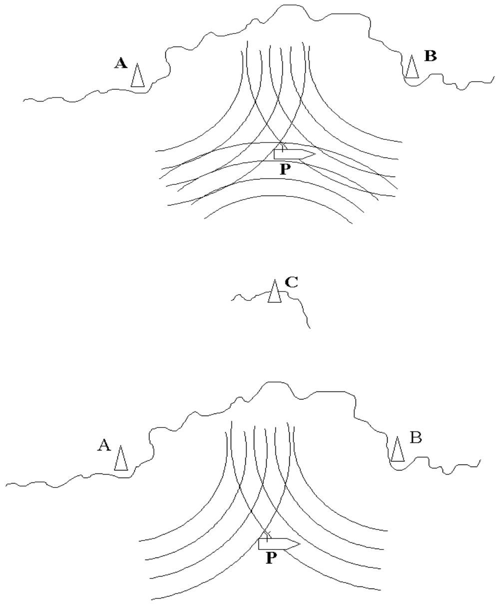


FIGURE 7.4

Circular Navigation in the Plane

In the case of space navigation, the position is obtained by suitable intersection of at least three spheres, whose centres and radii are known.

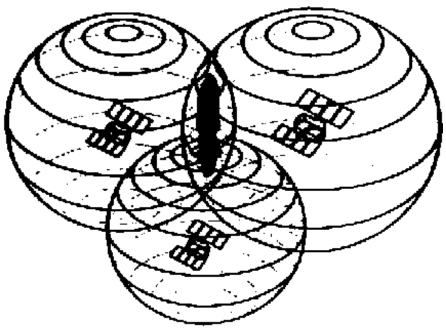


FIGURE 7.5

Spherical navigation example

Circular navigation can be of two sorts, named one-way and two-way. In the one way systems:

- all the stations and the users have precise clocks mutually synchronised.
- the time of arrival provides directly the time of propagation and consequently the distance.
- the number of users in unbounded (the user receives only).

Consequently, in principle, the implementation of the one way spherical method should require four FS, three on the spacecraft, the fourth in the unknown point, plus time interval measurement systems and time synchronisation systems.

It must be stressed again that all the four clocks (three on the spacecraft, the fourth on the user craft) must be of the same quality and accuracy and that they must be mutually synchronised. Such an implementation cannot be proposed for cost and complication reasons.

A very notable exception to the above mentioned requirements is offered by the satellite navigation systems, GPS and GLONASS, in which, by the use of a fourth satellite, fitted with an accurate and synchronised FS, the requirements of the clocks to be used on the user craft can be relaxed by 5-6 orders of magnitude. It is sufficient that the receiver contains a simple quartz oscillator with a stability of about 10^{-6} over the measurement cycle, which is typically of the order of seconds.

The NAVSTAR/GPS (Global Positioning System) is a space-born navigation and timing satellite system, operated by the USA Department of Defence. The spatial segment consists of a constellation of 24 satellites at 20 000 km height, distributed over 6 orbital planes and completing two orbits in a sidereal day. The system was designed to provide always at least six satellites on view anywhere on the earth. The radio signals are transmitted by the satellites on two L-band frequencies: L1 = 1 575.42 MHz and L2 = 1 227.6 MHz and the modulation signals are of the Direct Sequence Spread Spectrum (DSSS) type. They carry two codes: the "P" code, providing access to the Precise Positioning Service (PPS) and the C/A (Coarse Acquisition) code, used to access the Standard Positioning Service (SPS) that is quicker to lock on but is less accurate.

Four satellites are required for a navigation solution; pseudo-ranges to the satellites are computed by the user receiver from the measured time of arrival of the signals carrying the navigation message. This message contains, among others, the satellite ephemeris and its clock error versus the GPS system time. The basic equation used is $R = c \cdot T$, where R is the distance between each satellite and the receiver, c is the defined value of the velocity of light and T is the time difference between the satellites clocks (all ideally mutually synchronised) and the time of arrival of the GPS signals read in the user time scale. As the measurement of the satellites ranges is made by the internal clock of the GPS receiver, that is not accurate but remains stable for the measurement time, what is determined is a pseudo-range because it contains a bias due to the user clock error. Provided that pseudo-ranges from four satellites in suitable configuration are measured, a unique spatial fix in space is obtained by the user. For navigation purposes, one nanosecond of time error is equivalent to about 0.3 m of range error, therefore accurate synchronisation of the clocks onboard are extremely important in the GPS system. For this reason, the knowledge of the GPS system time versus UTC(USNO) time scale is maintained within 100 nanoseconds, and the corrections are made available to the users seeking for the ultimate accuracy.

The geographical coordinates plus the height, obtained from the spatial fix, are given by the receiver in the WGS – 84 geodetic reference system. The precision of the position obtained can range from tens of metres to a few centimetres, depending on the complexity of the receiver. For a detailed description of the system see ref. [Institute of Navigation, 1980].

An enhanced tool for real time precise navigation is now represented by the Differential GPS (DGPS), consisting in receiving, via an auxiliary communication channel, the information about the range errors found in the transmitted ephemeris of the visible satellites as computed by a receiver placed in a known position. The movable receiver can therefore correct the measurement performed and improve the position determination.

The GLONASS constellation, operated by Russia, when completed will also rely on 24 satellites with slightly different features. A detailed comparison between these two navigation systems is given in [Ponsonby, 1995].

In the two way systems:

- no synchronised clocks are required, the user ranges the reference stations, sending back and forth an electromagnetic pulse, the total time of propagation, divided by 2, times the velocity of light provides directly the radius of the circumference of the sphere,

- the number of users is limited (saturation of the telecommunication channels).

Consequently in the implementation of the two way spherical method only one FS is needed at the user side as a reference for its time interval measurement system. Two way circular navigation is performed by some precise positioning systems on ground, such as the Motorola Minitrack, and in space for deep space navigation.

The one-way, circular method, in its version with four satellites available in every point at the ground, at the moment and for the foreseeable future is the most important and widespread.

7.4.3 Hyperbolic navigation

Let us assume that, by suitable measurements made on a unique satellite, at a fixed ground-station of unknown position, a set of ranges r_1, r_2, r_3, ..., made in different times, for instance, every two minutes during a pass of the satellite, at UTC times t_1, t_2, t_3, are obtained. The satellite transmits the UTC time and its orbital parameters. Combining this last information, the positions of the satellite P_1, P_2, ..., etc. can be calculated at the given UTC times t_1, t_2, etc. Taking the differences $(r_2 - r_1)$, $(r_3 - r_2)$, etc., a number of hyperboloids are obtained, having foci at the positions of the spacecraft and parameters given by the range differences.

7.4.4 Hyperbolae, hyperboloids and their properties

It is recalled that, in the plane, the hyperbola is the curve whose points have a constant difference between the distances from two fixed points, called foci (Figure 7.6). In hyperbolic navigation in the plane, a receiver onboard measures the difference of the propagation time from two stations, located at the foci, and its current position. The propagation time multiplied by the velocity of light is the distance; with subsequent measurements, the difference of distances (transmitting station position minus receiving station position) is obtained. In such a way, one hyperbola is defined (continuous line in Figure 7.6). Repeating the measurement with two other stations, a new hyperbola is obtained (dotted line) and the position is given by the intersection of the two curves. In space (see Figure 7.7), the time difference measurements, give place to the hyperboloids.

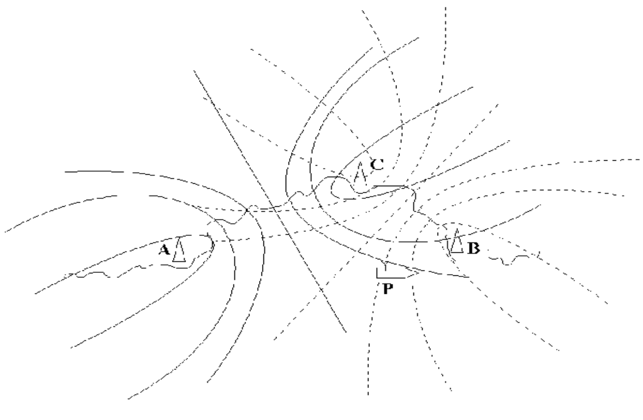


FIGURE 7.6

Hyperbolic Navigation in the Plane

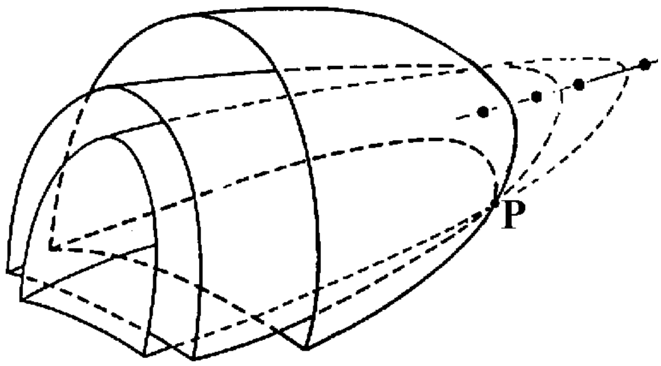


FIGURE 7.7

Hyperbolic Navigation in Space

Hyperbolic navigation in space can be performed using the signals provided by the satellite navigation systems TRANSIT and TSIKADA.

On the surface of the earth, hyperbolic navigation is provided by a number of systems, the most important being LORAN-C and OMEGA. In the implementation of the hyperbolic method, each satellite or ground station must be provided by an atomic FS; in some cases it is sufficient a source of frequency, in other cases also synchronised clocks are needed. At the user side, a simple FS is sufficient. The number of users is unbounded (reception only is used).

7.4.5 Accuracy requirements for the frequency standards used in navigation systems

The performances of the clocks and/or the frequency standards to be used in the stations or satellites providing -the infrastructure, as a rule the best state of the art devices i.e. atomic standards must be used.

On board the mobile platform, the accuracy of the clock and the frequency error and instability of the frequency standard depend on the method adopted and in the position accuracy that is needed. Some evaluations are made in the followings; in a pure one way circular navigation (precise clock on board), if the error in position after one day of navigation must be maintained within about 0.5 km, the setting error of the on-board clock at the beginning of the trip must be less than one microsecond and the maximum relative frequency difference between all the clocks involved, those of the infrastructure plus the on board clock, must be less than 2×10^{-11}.

In the two way method, no accurate clock is required on board or at the station; in one of the terminals of the link a simple transponder is needed and a time interval counter plus a radio equipment is used at the other. The type of the counter depends again on the accuracy needed in the range measurement. The accuracy ranges from 10^{-5} for ground applications, to 10^{-10} for space uses such as satellite laser ranging. The reader interested in details about the different navigation systems is referred to the standard reference books.

7.5 Telecommunications

Telecommunications at their very beginning (circa 1850) were, in principle, a low rate digital system (telegraph using the Morse code, presence or absence of current), but the bulk of telecommunication methods quickly followed an analogue approach, with telephone, phonograph, radio-broadcasting, radio and microwave links, magnetic recording, television, etc., at least for a century.

After 1970, a strong drive towards digital techniques has been observed; in the new telephony systems the three basic functions – the *coding and decoding* of the message, the *signalling* (connection from the calling subscriber to the wanted one) and the <u>*transmission*</u> are performed using digital techniques. Digital communications usually have requirements, regarding the FS used, far stricter than analogue systems. The latter are still extensively used and consequently it is of order to analyse the two cases.

7.5.1 Analogue systems

No very stringent requirements are needed in HF communications, the needed accuracy being in the region between 10^{-5} and 10^{-7}. A special case is represented by the frequency translation chains used in the microwave radio links, in which the base-board audio channel (300-3400 Hz) is transferred back and forth, in 6-8 steps, to the microwave carriers that link the repeaters. The ITU-T (former CCITT) regulations require than in a fictitious link of 2 500 km, the total frequency error in the base band should remain less than 2 Hz. In these frequency translation chains, the basic operations performed are usually <u>*sums*</u> at the transmitting side and <u>*differences*</u> at the receiving side. It is well known that, in sum and difference operations, the uncertainty of the result is the sum of the uncertainty of the two terms. The total frequency error of 2 Hz is consequently the limit of the sum of the relative frequency differences of a dozen or more of individual oscillators.

This limit requires that, in the chain of 6-8 quartz oscillators needed for the above mentioned frequency translation, the uncertainties range from 10^{-6} for the first steps to $10^{-8} \rightarrow 10^{-9}$ for the last steps of the frequency translation. Simple but effective means are used in order to achieve this goal.

7.5.2 Digital systems

The case of digital communications is quite different since the three above mentioned functions, coding and decoding, signalling and transmission, are performed by time-ordered pulses. In the digital system the *position* of each individual pulse inside a stream of similar pulses leads to the meaning of the symbol to be transmitted. The position in time calls for the first use of FS techniques (namely frequency standards, counters, etc.). The position inside the stream, called frame, requires in turn that the *beginning of each stream* should be identified unambiguously. This identification calls for a second use of FS techniques (namely frequency standards, clocks, counters and synchronisation systems).

It is evident that if there exists a frequency difference between the coding and decoding clocks, a phase difference grows between the two clocks. When this difference reaches the order of the length (in time) of a frame, the order of the pulse inside the frame cannot be identified anymore, a decoding error takes place and the meaning of the frame is lost. Digital communications are consequently deeply rooted in FS devices and methods. A number of different approaches took or are taking place, the most important two being outlined in what follows.

In the first approach, used in the most industrialised countries, the network is based on a hierarchy of clocks distributed in a number of levels called *strata*. Starting from the upper level, in which accurate master clocks are used, one finds, descending the pyramidal hierarchy, clocks of lesser and lesser quality. Synchronisation links between one stratum and the following one, have the character of a master-slave relation. This TOP-BOTTOM link has recognised merits, but problems are sometimes met in the signalling function. Inside geographically large countries, where more independent networks can be found, and in the case of the networks of different countries, each network is driven by a separate master clock or by an ensemble of separate clocks (all atomic) forming the upper stratum.

This approach, based in independent clocks in near synchronism, is called Plesiochronous Digital Hierarchy (PDH) and is the existing transmission infrastructure used in the international links. Plesiochronous means “close (Plesio) – clocks (chronos)”, i.e., clocks close, or near, in frequency, i.e., nearly synchronous.

The new approach, now in a phase of advanced study and beginning to be implemented, uses a more flexible structure of clocks, and is called Synchronous Digital Hierarchy (SDH) and will replace PDH. In the SDH a chain of spatially remote low quality clocks is inserted between two high quality clocks slaved to each other. The system is thus entirely synchronous. SDH provides easier digital signal multiplexing, signalling and cross-connecting. The network efficiency and robustness are improved. To allow for a smooth transition from PDH to SDH, the new standard must provide full compatibility with PDH. Less industrialised countries tend to skip the PDH standard phase and to pass directly to SDH. Technical details about the implementation of these two standards can be found on ITU-T (former CCITT) Study Group 13, ANSI and ETSI publications; here only the frequency and time requirements are covered.

For plesiochronous systems and considering the upper level, used in international links, the ITU permits (Recommendations G 811-G823) one decoding error every 15 days; the relation giving the total integrated error TIE (the phase error accumulated between the two interested clocks) is:

$$TIE \ < \ (10 \cdot s \ + \ 2\,500) \text{ ns}$$

where s is the duration, in seconds, in which no decoding error is admitted. The duration, s, depends on the service (voice, teletype, fax, data) and from the error rate accepted in a given link. The Bit Error Rate (BER) is indeed one of the most important parameters in designing the implementation of a network.\

For $s = 8.64 \cdot 10^5$ s (ten days), it turns out that TIE < 11 140 ns. Converting this phase accumulation in terms of relative frequency difference between the two clocks, one finds that $\Delta f / f = 1.29 \times 10^{-11}$. This figure calls for the use of atomic frequency standards plus a very complex synchronisation network inside a hierarchy of frequency standards located in the telecommunication centres of a national network. This synchronisation can be provided by external means (Loran-C, GPS, etc.) or performed inside the network.

To characterise the network performance in SDH systems, two parameters are used, the Maximum Time Interval Error (MTIE) (τ) and the TIErms(τ), both as a function of the measurement time τ. The standards prescribe for them the following limits:

$$MTIE(\tau) \ < \ 18\,000 \text{ ns with } \ \tau \ < \ 100\,000 \text{ s, and}$$

$$TIE_{rms}(\tau) \ < \ 300 \text{ ns with } \ \tau \ > \ 100 \text{ s}$$

The long term stability needed can be obtained by means of atomic clocks (rubidium, cæsium) or by synchronisation techniques using master-slave configurations. The short term stability is provided by quartz clocks, locked with special techniques.

The most appropriate balance between atomic clocks (higher cost, lower drifts, simpler synchronisation procedures, fewer decoding errors) versus a network formed mostly of quartz oscillators (lower cost, extended and proven reliability, higher drifts, complex synchronisation procedures, higher error rate) is a matter of debate. It can nevertheless be safely assumed that the problem lies in the network efficiency and capacity; if the error rate is significant, the computers surveying the traffic flow ask that a specific message is repeated back and forth until satisfactorily results are obtained. This procedure, if too frequent, lowers indeed the efficient use of the channel. It has been estimated that a 1% increase in the whole cost of a network, due to better clocks plus additional circuits, can increase of about 10% the capacity of the network.

The development of digital communications, at least from the point of view of FS, will require some progresses in FS devices, such as:

- the improvement of the reliability of atomic frequency standards, with a goal of about 10^5 hours of uninterrupted service,
- the production of small inexpensive standards with 10^{-10} - 10^{-11} accuracies for their life, to be used in small autonomous systems,
- the improvement of the spectral purity of the frequency standards since higher bit rates will require a high phase-coherence level also in the optical region.
- finally, a better understanding between producers and users of FS in telecommunications of the different terms, specifications and symbols used and a standardisation of the statistics will be mandatory.

7.6 Other applications

The low cost and the level of accuracy easily achievable with FS devices open new and wide fields of application, but the use of precise time and frequency references today pervades technology and science in such a way that any review is far from being exhaustive and some choices have to be made. In our case, only a few examples of applications will be given.

7.6.1 Automotive applications

Depending on the class of the car or truck, the number of quartz oscillators inside a vehicle ranges from a few units to a few dozens. In the following, some of their applications are listed; in addition to these, it must be remembered that each microprocessor used in a car hosts a quartz-stabilised oscillator.

Quartz units in automotive applications:

- clock-radio receiver, fitted usually with an indirect frequency synthesiser,
- engine,
- emission control,
- transmission,
- anti-skid microprocessors,
- timing sources,
- bus interfaces used for the electric power distribution,
- “intelligent” cruise control,
- lane access,
- anti-collision,
- automatic debiting,
- traffic information, etc.

7.6.2 Electrical power systems and compressed gas dispatching

When a large amount of energy is transmitted along power lines or pipelines, in order to ensure an ordinate flow of energy along the network, a number of time ordered measurements must be performed. The goal is to reach a power flow efficiency by way of a suitable power dispatching. In the case of electrical power systems, to help in the analysis of the system operation, the record of the perturbations occurred must be time-tagged with errors of less than 1 ms, over continental areas.

Moreover a smooth management of high power generators such as large turbo gas or large steam turbines require a strict timing. Electrical energy (alternating current) cannot be stored, and in the case of malfunctions, related current or voltage surges propagate along the lines. These surges are some times large enough to generate destructive effects on the line equipment and consequently long lasting interruptions in the service provided. In tracing the origin of the problem and in finding the appropriate measures, the time tagging of the individual events is mandatory. If a time synchronisation at the microsecond level, which is the present goal, is achieved, the fault along the line can be located with an uncertainty of 300 m, approximately the distance between power line towers.

A number of solutions have been developed over the years, based on time signals and standard frequency emissions radiated via dedicated or radio broadcasting networks and, more recently, on the use of space-based satellite navigation systems such as GOES and GPS [Wilson, 1991].

7.6.3 Instrumentation

A large number of electronic measurement instruments are based on FS. A few examples will be given below.

In the first case, the FS itself is used directly and is mandatory. The most relevant instance is an instrument called “electronic counter,” which is available in two basic versions: the frequency counter and the time interval (or period) counter. In the first case, Figure 7.8, the pulses of the FS are used to generate a time interval of known length T_s, *e.g.*, one second, that opens a gate.

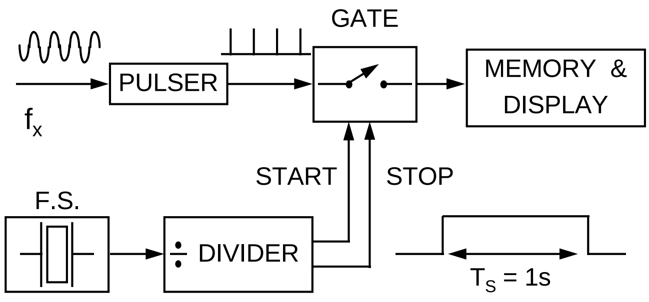


FIGURE 7.8

Principle of the frequency counter

The pulses of an external signal F_x, whose frequency is unknown, occurring in the interval T_s are counted; the result of such operation provides the frequency, F_x, in hertz.

In the second case, Figure 7.9, the number of pulses of FS falling inside the time interval defined by two events, occurring at to inputs named “start” and “stop,” gives the time-length of that interval, expressed in units of the period of the frequency generated by the FS. That interval can be defined by two consecutive zero crossings, in the same direction, of the input signal; in this latter instance the instrument is called period counter.

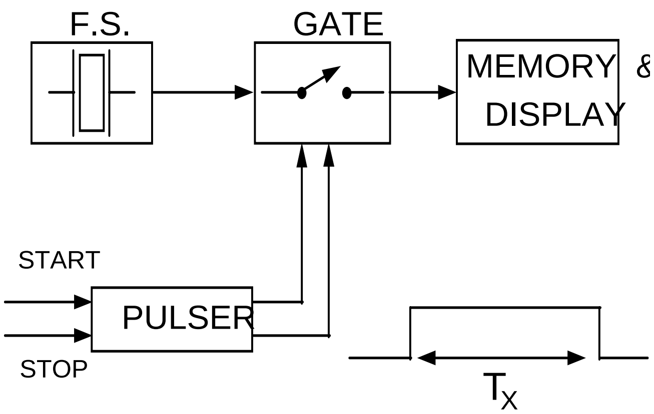


FIGURE 7.9

Principle of the time interval counter

Let us assume that the nominal frequency of the FS is 100 MHz, and, for example, the “time distance” between the start and stop pulses is exactly 1 450 ms; in this case the final content of the memory is 145 000, in units of 10 ns, i.e. 1 450.00 microseconds or 0.001 450 00 s, the instrument allowing directly the choice of the unit used (s, ms, µs, …), and consequently the presentation of the result in the most suitable way for the user.

These two basic approaches have found a number of different implementations along the years. In our second example, the FS, used as a frequency reference from which the output signal is derived; brings a substantial improvement to the instrument capabilities. The instrument is the synthesiser, a device driven by one FS and producing a very large number of signals over a large frequency range, each one supplied with the accuracy of the FS. One commercial instrument, for instance, can produce 5×10^{10} individual signals in the range from 1 Hz to 50 MHz, with a minimum step, from one signal to the next one, of 0.001 Hz. Frequency synthesisers are widely used in research and telecommunications.

In the third case, a FS is used in an instrument that measures any quantity other than frequency or time, because its use provides accuracy advantages or simplicity in design and construction or eventually an easiness of use. This trend is particularly marked in instruments providing a digital output. As an example, many digital voltmeters convert the unknown voltage into a frequency or into a time interval that is measured at last with an electronic counter.

7.6.4 Doppler radar

In many instances, the quantity of interest is the relative velocity between two bodies in motion or between a body in motion and one at rest. This measurement is usually performed via the Doppler effect, previously introduced in Section 7.4 concerning some navigation methods. Doppler radars are extensively used in aeronautics and are proposed for the car collision avoidance systems. In these devices, the relative velocity is transformed in a frequency variation that is eventually measured with an electronic counter. Doppler radars find use also in other unexpected applications, such as the intrusion detection or the docking of very large crude carriers or supertankers. These ships can reach a mass of 5×10^{8} kg and must approach the dock with a controlled speed of about one millimetre per second. Measuring such speeds with a Doppler radar, when the round trip propagation time is less than 100 ns, poses indeed very stringent requirements as regards the short term frequency instability in the radar oscillator on ground.

References

PETLEY B.W. [1988] The Fundamental Physical Constants and the Frontier of Measurement. Adam Hilger, London.

PÖPEL R. [1992] The Josephson Effect and the Voltage standard. *Metrologia,* n.29, pp. 153-174.

HARTLAND A. [1992] The Quantum Hall effect and Resistance standard. *Metrologia,* n.29, pp. 175-190.

INSTITUTE OF NAVIGATION [1980] Global Positioning System. Navigation, Vol. I, Washington D.C. (USA).

PONSONBY J.E.E. [1995] Global Satellite Navigation Systems – Uses of Space – Time fixes from Geodesy to Sailing. U.R.S.I. Space and Radio Science Symposium, Brussels (Belgium).

WILSON R.E. [1991] Uses of precise time and frequency in power systems. *Proc. of the IEEE*, Vol.79, n.7, 1009-1018.

Bibliography

METROLOGIA [1995] Special Issue on physical units, Vol.31, n.6, February.

FORSSELL B. [1991] Radionavigation Systems, Prentice Hall International Ltd.

KARTASCHOFF P. [1991] Synchronisation in Digital Communication networks. *Proc. of IEEE*, Vol.79, n.7, 1019-1028.

Chapter 8

Operational experience, problems, pitfalls

Contents

8.1 Frequency and Time Tools

A frequency and/or time system is generally composed of several parts depending upon the needs of the system. The heart of the system is the source of frequency (see Chapter 1 and 2). The cycles of this source are counted to keep time, which makes it a clock. Often an ensemble of units is used for reliability reasons and for the detection of errors or failures and to access the performance of each frequency reference and/or clock. Confidence in the performance of the source is the immediate question. Internal and external comparisons bring an answer which needs a judicious interpretation. Storage capabilities of the frequency/time units themselves, as well as the data being generated by them, are necessary in order to analyse the data, date the events, detect the errors and improve the frequency/time source. Often continuous operation is needed, then fail-safe power supplies will be needed to sustain this operation. How will the signals be disseminated for use. How much room is needed for all of the equipment?

Making a judicious choice in frequency/time equipment for a specific application is not a trivial job. National and international timing centres and telecommunications organisations are good places to start for helping the user in this job; in particular, they have a large operational experience which can be of interest.

In this section we consider issues like reliability, accuracy, stability, traceability and how to characterise frequency/time (F/T) systems. In the next section, data and examples are presented.

Prime requirements of the users of F/T refer to the availability of sources working properly, *i.e.*, in agreement with the official definitions and the known national and international references which are realised. As soon as the requested characteristics (stability, accuracy) are beyond those guaranteed by the manufacturer, the user faces various questions such as: How to be sure a F/T source works properly? What are its metrological characteristics? How to avoid any time discontinuity? How to manage the system in a cost-effective way? How to maintain and assure the performance of F/T sources?

8.1.1 Choice of a reference

Given the wide variety of uses of F/T devices and equipment, the choices vary dramatically. It is best to first address the needs and requirements along with the resources before acquiring equipment. As mentioned above, it is also best to consult with those who have had experience in a given area. Also, significant progress has been made in recent times in F/T sources and equipment. The advent of precise F/T signals from satellites has had a major impact on the choice of F/T reference sources. High accuracy in F/T can now be obtained at very low cost.

In the past, F/T systems have tended to be complicated and dependent on the available technology. Now, there are large systems requirements for precision F/T, and these needs are driving the designs. Because of the large numbers needed, the costs have dropped substantially in recent times. A much broader selection of excellent F/T sources are now available than heretofore.

8.1.1.1 Assessing what is needed

What is the need for a frequency reference and/or time reference? Does it need to be traceable to the Sl second and/or to UTC? Does it need to be kept within some boundary frequency limit or time limit as part of a network? Is it going to be compared with another source of F/T, and if so what are the stability and accuracy characteristics of that source and its comparison link? Each of these needs can require very different equipment and resources. Prudent design of the reference to meet the needs and requirements will be one of the most cost-effective things you can do. Many systems have been improperly designed because these questions were not asked and properly addressed, and large sums of money have been unnecessarily spent.

As an example, the ITU-T telecommunication frequency specification of 1×10^{-11} doesn't require time – only frequency (syntonisation). This specification arose in an era when it was much easier to obtain syntonisation than synchronisation. Because of satellite timing systems, synchronisation is now much easier to obtain than before. Many of the new F/T devices take advantage of this fact: a properly designed reference oscillator coupled to a synchronised system also yields syntonisation. With the development of SONET and SDH and for efficiency and fault-finding in power distribution systems, some of these new inexpensive capabilities are becoming increasingly important.

For standards laboratories, the choice of a reference is of prime importance in time and frequency metrology. Since one cannot measure the time of a clock, but only the time difference between two timing sources, these time-difference data are stored. These data are indications of time (or phase) differences between two clocks, or between one clock and a paper time scale (as defined in Chapter 6), or between two time scales. The time differences can be referred to as measurements characterising the combined variability of the two clocks or time scales. Sometimes the reference clock or time scale is considered perfect and all the variability is ascribed to the other clock or time scale; this, of course, is not reality.

Frequency data, on the other hand, can be traced to fundamental physics through the definition of the **Sl** second. Often, however, frequency measurements are stated in relative values between two standards or between one standard and an agreed upon reference. In contrast, official time should be referenced to the artefact standard, UTC.

8.1.1.2 Steered versus free-running clock

When a laboratory has several F/T sources with similar characteristics (*e.g.*, several cæsium standards of the same type), it is usual and, from a practical point of view, necessary to choose one of them as reference or master clock. Its choice will take into account the metrological official characteristics, the experimental behaviour with respect to the other clocks and the probability it will run without failure. Determining the best one from a performance point of view can be done using the n-cornered hat procedure or using a relatively new and more efficient procedure developed by Lepek [Lepek and Walls, 1993]

In the long term, the reference clock, regardless of its quality, will always tend to a random-walk in frequency – if it is free-running – from any other defined time scale. Hence, the options are to adjust the clock (steer its frequency to control the time) or to have a synthesiser (independently steered) on the clock's output. The first option effectively destroys the clock from being used as an independent contributor to a clock ensemble, and is not recommended. Some of the new standards have built in synthesisers so that one can have both an independent clock as well as a steered output from the same unit.

Beside the practical local reference, it is convenient, if possible, to have links with other references, national and international. There are two main advantages: first, they are external references which are generally considered as independent of the local one; second, they have recognised and quantified metrological characteristics in stability and accuracy. Results referring to such external references are often known only with delays. In the case of UTC, the delay is now more than one month.

A specific feature of the F/T metrology is that, with adequate equipment, any user is easily linked to the national references and, hence, to the international ones. So, in general, there is no difficulty to be in touch with several references which can contribute to the qualification and quantification of a source. In cases of isolated systems (submarines, spacecraft. . .) links are more difficult and may need to be intermittent. It is good practice to have more than one comparison link working at the same time. In this way, links can be compared as well as clocks.

In brief, it is important for the user dealing with F/T data to have in mind the concepts of redundancy (for the sources, for the references), of independence between the source and the reference, and the necessity of comparisons.

8.1.2 Tools for operational use

Among the tools which have been used to characterise the F/T equipment, two of them are of prime interest from the operational point of view. One aims at measuring the ability of a F/T source (or a time scale) to deliver the same time interval, namely to measure its frequency stability. A specific variance (two sample variance or Allan variance) has been commonly used; it is detailed in Chapter 3. Here, we will give a few remarks related to operational aspects of this variance.

The second tool refers to the notion of reliability. An in-depth inquiry was carried out, a couple of years ago, in the framework of the ITU-R (ex-CCIR) studies [CCIR, 1990]; the main values are shared in this Chapter along with some new materials, especially related to GPS receivers. There is no such information on the most recent standards. The user will need to seek help from the F/T centres and laboratories.

A third tool, namely the frequency accuracy, is often used. As explained in Chapter 1 and in the Glossary, it can relate specifically to the definition of the Sl second, which provides a fundamental reference rate for all clocks, or in a more general sense it can refer to other useful frequency standards. Unfortunately, in spite of the importance of this tool in F/T metrology, some misunderstandings have arisen; so caution should be exercised in its use.

8.1.2.1 F/T system stability

The stability of an F/T system will be limited by four things: the measurement noise, the instabilities in the contributing clocks, the processing noise and/or algorithms used to combine the readings, and instabilities in the F/T distribution. The final output will be the composite of all four contributions. An effort is usually made to make the measurement noise less than the clock noise. This can be accomplished in a well designed system and will be discussed further below.

In general, the frequency instability (or stability) is defined in Recommendation ITU-R TF.686 (see also Glossary) by "the spontaneous and/or environmentally caused frequency change within a given time interval" and it distinguishes between systematic effects and stochastic fluctuations. The latter fluctuations have been commonly characterised by the two sample variance (or Allan variance) for several years both by the users and the standards manufacturers. This variance has become an operational tool with numerous advantages.

- First of all, it is convergent for the usual frequency standard noises.
- Second, it is an unambiguous indicator of the type of noise which predominates in the frequency standard data with the exception of white and flicker phase noises which are not discriminated (see Chapter 3). In practice, this ambiguity problem usually only arises in measuring the short-term stability of quartz crystal oscillators and of hydrogen masers.
- Third, it deals with frequency drift terms (quadratic of phase data) in the same way as the random terms. This is easily explained by the fact that the two sample variance works with the phase second differences.
- Fourth, it has a signature corresponding to periodic phase variations.
- Fifth, it is easy to compute from sampled data. The simplest way is to take adjacent pairs of data. It has been shown [Stein, 1985] that the use of overlapping pairs leads to better estimates of the variance than adjacent ones [NIST, 1990].
- Sixth, through the three-cornered-hat approach applied to the variances of pairs of three independent frequency sources compared at the same dates, it is possible to estimate the stability of each individual unit. [NIST, 1990] A new approach can also be used with any convergent variance (including the Allan-variance) and obtain an even better estimate of the individual clock's stabilities than with the three or N-cornered hat approach. [Lepek and Walls, 1993]

From the operational point of view, the necessity of periodic data to compute the variance can be a difficulty. Smoothing of non-periodic data to get periodic ones can introduce insidious biases.

A modified Allan variance has been developed recently and has begun to be used by the metrology community. It has the advantage of discriminating white and flicker phase noises. The user has to be cautious that the modified two sample variance values are usually different from the two sample variance ones: they are typically smaller than the latter ones and the ratio depends on the type of noise and on the sample duration.

More recently a time variance (TVAR) has been developed. It has been readily adopted by the USA telecommunication community and is useful for characterising the random variations in measurement systems, in F/T distribution systems and in networks. It has most of the desirable properties of the modified Allan variance, but deals directly with time (or phase) residuals. (See Chapter 3.)

8.1.2.1.1 Measurement noise

An ideal time (or phase) difference measuring system will have white PM residuals. If processed optimally (using mod.$\sigma_y(\tau)$) , they will average as $\tau^{-3/2}$ (see Chapters 3 and 4). Systematic time delay variations cause measurement systems to depart from ideal, and the residuals may be better modelled by a flicker PM process. In this case they average as τ^{-1}. Most clocks (See Chapter 5) average as $\tau^{-1/2}$; hence, for long-enough averaging times the measurement noise can usually be made less than the clock noise. With improper system design, however, and with state-of-the-art clocks, this averaging time may need to be of the order of days. This is not necessary; with a proper design one can make the measurement noise be less than the clock noise for most values of τ. (See Chapter 4.)

Many timing centres use the 1 PPS from each clock to measure the time difference between it and the other clocks in the system. This is one example of improper design mentioned above. The 1 PPS is a fast-rise time pulse going into a high-frequency front-end time-interval counter. This design requires a very high-frequency measurement bandwidth, f_h ,which can adversely affect the measurements. The sophistication of the time-interval counter and of the 1 PPS generation must be state-of-the-art to push the effective measurement noise down. Even then, long averaging times are required, and the equipment costs are high.

A double heterodyne technique has been developed, which avoids many of the above problems. One version of it is referred to as a dual-mixer time difference technique. It avoids the 1 PPS divider noise by measuring the phase of the RF signal of the clock, and desirable and controllable narrow measurement bandwidths are easily obtainable, *e.g.* well less than 1 kHz. The measurement noise can be made less than the clock noise in most cases for averaging times as short as one second. (See Chapter 4).

8.1.2.1.2 Measuring clock performance

It is impossible to measure the performance of an individual clock by using only the measurement data against another clock of comparable quality. If one is significantly better, then we can approximate the performance of the worst, but this cannot be determined by only having measurements between the two clocks.

At least three or more clocks are required to estimate truly the performance of each of the clocks in the set. If the N-cornered hat approach is used, negative variances will arise from time to time, and there is always the question of what to do with them [NIST, 1990]. A relatively new approach avoids this problem while obtaining a better estimate of performance. [Lepek et al., 1993]

Using clocks in an ensemble opens an opportunity to assess the individual performance of each of the contributing members. Even though each clock will be somewhat correlated with the ensemble time, since it contributes to the value, this correlation can be accounted for, and an unbiased estimate of the performance of each member clock can be determined. (See Chapter 6.)

8.1.2.2 System reliability

In many modern applications, reliability is the most important item. A system should be designed to avoid single points of failure or to make near zero the probability of a catastrophic failure. Ensembles and system redundancy are ways to increase reliability if proper design criteria are used. Care is very important here, because more complex systems can be less reliable.

8.1.2.2.1 Failure rates

The traditional way of measuring the reliability of a device is through the mean time between failures (MTBF) statistic. It is estimated by dividing the period of running (in agreement with the specified characteristics) of an ensemble of similar devices by the number of failures during the period of time; units can be years, months, etc. The estimate which is obtained is meaningful only if the reliability of the device can be considered as constant with respect to time. The MTBF measure has the significant disadvantage that one has to wait until every unit fails before the mean can be computed.

A better statistic for reporting reliability is the probability that a unit, having survived a time t, will fail by time $t + \Delta t$. It has been defined as a conditional failure rate $Z(t) = -\Delta N(t) / N(t) \Delta t$ where $N(t)$ is the number of devices running at t, $\Delta N(t)$ is the change in this number from t to $t + \Delta t$.

Estimates of $Z(t)$ are obtained by dividing the number of devices which have failed during a specific period (one year, for example) by the total time of running (counted in years per the example) of the devices.

Computation of $Z(t)$ was carried out in January 1970 at the United States Naval Observatory for an ensemble of cæsium clocks and led to $Z(t) = kt$ with $k = 0.1$ for $t > 1$ year.

From the $Z(t)$ model, it is possible to compute the half-life (HL) [Percival, 1975]; the probability that a clock survives to a half-life time is 50% and the estimate of HL is available after one half of a test set of clocks have failed.

8.1.2.2.2 Problems in dealing with errors

Significant confusion often arises in dealing with error specification. What kind of variance is used to measure the errors? Is the error statement a 1σ, 2σ, or a 3σ value, or is it based on extremes over some running time? Is the distribution of errors normal (Gaussian)?

It is important to follow good procedure and to clearly specify what has been measured. Clock error distributions have been typically found to be close to Gaussian in their distribution, but often there will be more energy in the wings than for a normal distribution.

8.1.2.3 System accuracy

Accuracy in frequency is defined with respect to the SI second. Accuracy in time is defined with respect to UTC. From a systems point of view, self consistency of the time and/or frequency may be sufficient for successful operation.

If you buy a frequency source which is supposed to deliver, after a normal warm-up period, a 5 MHz signal and you compare it, during a certain interval (hours, days) with a frequency reference and you obtain 4.95 MHz; you announce the accuracy (or the inaccuracy) is 1% (i.e., nominal offset measured over nominal). In common sense, the wording sounds correct. In a metrological sense, it does not. What has been expressed by 1% is one normalised frequency deviation or difference measured over one specific time interval around one date. There are three potential problems: how accurate is the referenced used with respect to the SI; how much uncertainty is contributed by the measurement; and, since it is only one measurement at a certain time, how do you know it will not change as time goes on.

The above does not express an accuracy as defined in Recommendation ITU-R TF.686, “degree of conformity of a measured or calculated value to its definition (see uncertainty)” and uncertainty, “limit of the confidence interval of a measured or calculated quantity”.

Clearly, the metrological accuracy concept has to be separate from the general common sense accuracy approach. Any measurement aims at linking a quantity of interest with the corresponding unit, the link is expressed by a number and a unit. It is a generally desired that the number quantifying the measurement be as correct as possible and reflect, as much as possible, the true value. In other words, any measurement responds to a general accuracy approach.

In frequency/time metrology, the frequency accuracy concept refers to the qualification of the general behaviour of a standard, so, it is valid on the very long term: for the life-time of a cæsium tube in commercial standards, for the period where a laboratory standard (and the measurement procedures) is maintained unchanged. Furthermore, according to the ITU-R definition, the accuracy number expresses the conformity between the measurements of the quantity and its definition. Strict application of that limits the use of accuracy to cæsium standards working in agreement with the definition of the second, or some way of providing traceability to the definition of the second. Finally, accuracy is expressed by an uncertainty value. In laboratory cæsium standards, all the possible sources of uncertainties attached to the running of the standard are taken into account and the final estimate results from random terms analysis and from statistical interpretation of possibly deterministic terms (Chapter 1).

It is usual in frequency/time metrology to broaden the use of the accuracy concept to non-caesium standards, such as H-maser, rubidium, lasers, and also to systems used to compare clocks, such as LORAN-C, GPS… In the latter case, time accuracy may also be an issue.

8.1.2.4 Practical hardware problems

This section, of course, could be a very long list. Given limited space, we only list here a few of the most common ones, which are often overlooked.

Significant disparities often occur in dealing with time accuracy. When working at the nanosecond level, this requires all delays to be known to better than the accuracy goal. Electrical lengths of cables, antenna feeds, amplifier delays, etc. need to be known. Where symmetric cables are used, for example at the front end of a time interval counter, their electrical lengths need to be the same at the decimetre level and be properly terminated. The 1 PPS rise-times and trigger points become very important and there is no standardisation in this regard. Some labs trigger at 0.4 volts, some at 0.5 and some at 1. In defining a time scale, there needs to be a defined phase plane at which point time, the matching impedance and type of connector, for example, should be specified.

- **Ground loops**

  When multiple instruments are interconnected, there is a problem of having multiple ground connections through which rather large currents can flow because of the low loop resistance. These currents cause extraneous voltages at the inputs and outputs of the various instruments. These voltages can degrade the performance of the Frequency Source, the distribution system , and the measurement instrumentation. The best solution to this problem is to have all the ground connections made at a single point. This may require isolation transformers. Any text book on the use of low-noise operational amplifiers can give useful considerations on eliminating this problem.

- **Power-line related magnetic fields**

  Electronic instruments contain power supplies which often have stray magnetic fields at the power line frequency (50 or 60 Hz). Depending on the quality of the supplies and the position in the electronics rack, this field has been measured to be as great as 50 µT (50 milligauss). Careful positioning of the various instruments or magnetic shielding may be necessary to reduce the field enough so that it does not degrade the performance of the frequency source, see Chapter 5.

- **Temperature control**

  Ordinary room temperature control may not be sufficient to allow the frequency source to perform at its capability. The temperature variation is typically ± 3° C or greater and, with on-off (bang-bang) types of controllers it has a periodicity between 2 minutes to 20 minutes. If the frequency source is sensitive enough, this may produce a periodic variation in the output frequency and degradation of the power spectral density of phase at a frequency corresponding to the period of the air conditioning system. High quality air conditioning with proportional control can reduce the temperature variation to ± 0.05° C for a room. If a less expensive solution is required, some small temperature controlled boxes are available. Egg hatching incubators can work very well.

- **Vibration**

  Vibration is usually present in a laboratory, caused by air conditioning systems and instrument fans. The levels are typically at a level of 0.2 m/s^2 at a frequency which depends on the rotation period of the fans and motors. These frequencies are high enough that some vibration isolators can reduce the effect. It may be necessary to mount the frequency source on a foundation which isolates it from the building vibration. The vibration effects can be seen on a low-frequency (Fast Fourier Transform) spectrum analyser.

- **Power line stability**

  Some frequency sources are sensitive to power line voltage changes. In addition, after a power outage, the frequency source may require several days or longer to reach a stable, accurate frequency again. For these reasons it is advisable to operate the frequency source from an uninterruptible power supply (UPS). A number of good commercial units are available. It is important that the size of the battery in the UPS be large enough to last for the longest expected outage and that the batteries be changed periodically since they have a limited lifetime. For frequency sources that are designed to operate from a back-up battery system, that may be the only UPS required.

- **VSWR, cables, connectors**

  The commonly used BNC connectors are notorious for mis-match and for delay variations and step changes. As an example of a cable length problem, the wave length at 5 MHz is about 60 m. Unterminated cables with 1/4 λ (15 m) electrical length, which is not uncommon for general use, can act like a short. Cables themselves can cause problems because of phase changes and jumps when they are moved or when the temperature is changed. Cables such as RG-58 cause the most trouble, followed in order by RG-223, then any of the phase table cables such as Flexco, and lastly the large, air-dielectric semi-rigid cables. Similarly, the connectors, in increasing order of quality are BNC, TNC, SMA, and precision N. It is important that the VSWR on the cables be minimised by using the proper termination resistance. High VSWR reflects reactive impedance into the output of the frequency source and may cause a change in the frequency. Small changes in electrical length can then cause changes in the frequency which would not be present if the cables were properly terminated (see paragraph on isolation). It should be mentioned that the most stable cables are fibre optics cables, but a fibre optic system must be carefully engineered for frequency and/or time distribution to operate correctly.

- **Isolation of output amplifiers**

  If the output of the frequency standard is connected to another device with a poor VSWR or which exhibits leakage of internal signals, the frequency may be affected. The isolation of the output amplifiers is very important to keep external devices from changing the frequency of the frequency source. Some of the frequency sources do not have adequate isolation, so external high isolation amplifiers may be necessary. These are generally incorporated in a distribution system which has one input and a number of outputs. Isolation of 100 to 130 dB is adequate for even the best frequency source.

8.2 Data and examples from operational experience

8.2.1 Frequency and time standards

Stability and accuracy data for atomic frequency standards (rubidium, cæsium, hydrogen maser) are presented in Chapters 1, 2, and 5. They can be used as basic material to select appropriate frequency/time to satisfy user needs.

For many applications, quartz crystal standards (see Chapter 1) are appropriate, especially if short-term stability (times less than 1 s up to 10 s) is required. Furthermore, recent developments of quartz oscillators (or rubidium) steered by a GPS signal have resulted in frequency/time units which exhibit both the short-term stability of the quartz and the long-term stability derived from the cæsium standards of GPS. The European Frequency and Time Forum, the Intentional IEEE Frequency Control Symposium, and the Precise Time and Time Interval Planning and Applications meetings proceedings are excellent sources of information about these developments.

Reliability data are of importance, especially for those in charge of planning the maintenance of frequency units and to ensure the reliable long-term performance and the security of systems. MTBF has been estimated for cæsium, rubidium, and crystal clocks. Results relative to the mean time to repair (MTTR) a failed unit including shipping time have also been mentioned. These MTBF and MTTR values are given in Tables 8.1, 8.2 and 8.3.

TABLE 8.1

MTBF and MTTR values of cæsium clocks

Model (year)	$\sum$U	$\sum$F	MTBF (years)	MTTR (days)	No. of units in survey
Users' report					
HP5061A (1968)	3347	823	$4.07^{+0.69}_{-0.52}$	90	492
OSA 3200 (1975/76)	96	32	$3.0^{+0.6}_{-0.4}$	90	25
HP5061A-004 (1973)	118	44	$2.68^{+0.41}_{-0.31}$	90	24
HP5060A(1) (1965)	133	42	3.17	90	21
OSA 3000 (1976)	29	10	$2.9^{+0.9}_{-0.6}$	90	14
HP5062 (1973)	1648	319 (2)	$5.2^{+1.0}_{-0.7}$	90	408
Manufacturers' report (see comment § 3.3)					
OSA 3000 (1976)	285	30	9.6	35	97
OSA 3200 (1975)	679	161	4.22	50	149

(1) Old model, no longer in production.

(2) Only cæsium beam tube failures and other failures associated with beam tube failures are included for this particular type of clock.

TABLE 8.2

MTBF and MTTR values for rubidium clocks

Model	$\sum U$	$\sum F$	MTBF	MTTR	No. of units in survey
HP5065A (1970)	159	21	7.6	120	
FRT/FRK (I 973)	584	52	11.2	90	159
XSRM (1972)	71	13	5.5	90	15
POI (1976)	44	41	1.08	------	20

TABLE 8.3

MTBF for crystal clocks

Model	1st year	$\sum U$	$\sum F$	MTBF (years)	No. of units in survey	Notes
Users' report						
B5400	1974	48	1	48	11	
B1250	1973	8	1	8	1	([1])
B1010	1965	926	25	37	132	([1])
HP104/105	1970	46	4	11.5	5	([2])
R&S XSC/D/S	1970	136	13	10.5	15	([2])
C60MCS	1972	223	1	200	52	
CP12MCS	1970	6316	33	191	1288	
MT	1975	834	13	64	139	
K	1975	1353	2	200	235	
Manufacturers' report						
OSA B5400	1974	1352	27	50	318	
OSA B1250	1970	2314	3	71	20	([1])
HCD HCD50	1970	4383	104	42	587	

([1]) Obsolete: no longer manufactured

([2]) Units combined in a single survey because of high similarity of design and no apparent bias

MTBF = $\sum U / \sum F$, where

$\sum U$ is the sum over all units of the number of years of operation for each unit, and

$\sum F$ is the total of failures observed for these units.

Detailed analysis of the distribution of failures among the various sub-ensembles has been performed for cæsium standards: it appears that the atomic resonator (the cæsium tube) fails in about 27 to 40% of the cases. That is an important point since the price of a cæsium tube is about half the price of the standard. It is worth noting that the guarantee period of the cæsium standards has been increased by the manufacturers during the last years, going up to 5 years (or more) in some cases.

8.2.2 Examples of problems

Time and frequency metrologists try to get the best from the commercial standards running in their laboratories (especially cæsium standards and hydrogen masers) and they look after them with great attention. Here are examples of problems they have identified.

- Correlations have been reported between the phase signals of cæsium clocks operating near each other.
- Environmental conditions of clocks (temperature, humidity, magnetism,..) are considered as important for maintaining a good stability of the signals. In particular, humidity was often quoted as a possible agent of instability. It is worth noting that, in its yearly inquiry, the BIPM asks for data on changes in the environmental conditions of clocks.
- Air pollution was reported [Freon, 1990] as being at the origin of the deterioration of cæsium standards switches, and also of small phase jumps. In the same reference, electronic settings and adjustments were clearly identified as sources of frequency jumps (on the order of 10^{-13}) and phase jumps (up to tens of nanoseconds) in cæsium standards. These problems ought to be largely reduced by the introduction of data processing systems and digital techniques in the most recent cæsium standards of the 90's.
- A complete monitoring system is necessary in order maintain the integrity of the frequency sources. Archiving of data in a form which will enable one to examine long-term changes is a necessity.
- A repair/replace philosophy must be established to set up criteria which will enable a rational decision to be made as to whether it is time to repair or to replace a frequency source or not.

A few laboratories (VNIIFTRII in Russia, USNO in the USA, etc.) have extensive experience with large ensembles of hydrogen masers operating continuously. The cavity autotuning units of these standards is a crucial point, it can be the origin of discontinuities and of the degradation of the spectral density of phase [Audoin, 1992].

8.2.3 Frequency and time comparisons

As mentioned above, comparisons of frequency/time units are essential for the realisation of performing and reliable frequency/time systems. National and international comparisons are used by taking advantage of the large possibilities available from satellites; the GPS ones have been extensively used by the metrological community (BIPM) since 1983. More recently, the GLONASS has been investigated (see Chapter 2B).

Calibration of the comparison results is an important and difficult topic, especially in time metrology where it impacts directly on the qualities of the time references. It can be dealt with in two ways: first, a global approach by transportation of a clock (technique used up to the 1980's) or of a GPS (or GLONASS) receiver. The uncertainty of the calibration result can be as low as a few nanoseconds, depending essentially on the care of the experiment.

A second approach [Lewandowski and Thomas, 1991] goes through a budget of uncertainties, taking into account, as much as possible, all the sources of uncertainties (random and systematic aspects have to be considered). This list exhibits points referring to:

- the antenna (coordinates, environment, cable ...);
- the receiver (time delay, software ...);
- the satellites (ephemerides, corrections ...);
- the statistical treatment of data (number of passes, biases, noises ...).

Time calibrations and studies on GPS capabilities have shown the existence of correlation between the GPS results and temperature [Lewandowski and Tourde, 1990] with several types of receivers. Phase jumps up to tens of nanoseconds originating in the GPS receivers have also been observed on several occasions. Clearly, redundancy in GPS receivers and in GPS comparisons is necessary to get the ultimate results in time metrology. The problem is not so crucial in frequency comparisons in principle. Nevertheless frequency comparisons at the level of 10^{-15} or 10^{-16} with reasonable time lags asks for receivers more stable than are currently available.

8.2.3.l Is linear regression best?

If the residuals around the regression line have a white-noise spectrum, then the fit is best. This would typically be the case, for example, between two cæsium standards: since the random fluctuations are modelled by a white FM spectrum, the optimum frequency difference is the simple mean. Subtracting the first time residual from the last and dividing by the elapsed time is equivalent to the mean – where the mean is the normalised frequency difference between the two clocks. If this end point method is used, caution should be exercised since either one or both of the two points could be a statistical outlier.

If a frequency drift is being estimated between two cæsium standards, then the regression line to the frequency is optimum. If the linear regression line is fit to the time differences, then the slope of this line is far from an optimum estimate, because the residuals would typically be random-walk. The inefficient method of making a linear regression fit to the time (or phas*e)* difference to determine the time and frequency difference is all too common within the F/T community. If the long-term fluctuations between two clocks may be well modelled by random-walk FM, then the second difference is the optimum frequency drift estimator. This can efficiently be obtained from the first, middle and last time residual point [NIST, 1990].

A frequency offset is sometimes referred to as a time drift. This is not a good designation, since there will always be a frequency offset between any two clocks, *i.e.* frequency offsets are normal, while drift sounds adverse.

8.2.3.2 Problems with cycle ambiguity

Often it is efficient and convenient to use, for example, a zero-crossing of the RF phase as a time marker. The main problem with this technique is the resolution of which cycle is being measured. For example, at 5 MHz one cycle takes 200 ns to elapse. If the time is known to less than 1 radian of the phase, then the cycle can usually be resolved. Since the day-to-day RMS deviation between two cæsium clocks is usually less than 10 ns, cycle resolution is not usually a significant problem for these clocks. With Loran-C, however, where a cycle is 10 μs, cycle slips do occasionally occur. Over long baselines using LF and VLF systems, identifying cycles can sometimes be a problem because of mutual interference and ionospheric disturbances.

8.2.4 Other data, system set-up, and processing ideas and problems

Utilisation of and storing of clock time data in a well ordered way is critical to F/T metrology. Also, the utilisation of the data in optimum ways can save time and increase the effectiveness and the efficiency of the output. As examples, the NIST AT1 algorithm takes data every two hours at which point the time scale is computed for near real-time metrology purposes. In principle, such algorithms may produce an output which can be more stable than that of the best contributing clock in the clock ensemble. There is some risk in detecting both time and frequency adverse steps at a such short sampling rate but, none the less, this scale has operated continuously with only minor perturbations since 1968 and now has a flicker floor of 5×10^{-15}. The more conservative approach is taken by BIPM and USNO. USNO computes its official time UTC(USNO) a month after the fact – carefully accounting for any adverse behaviour in any of the contributing clocks. The USNO operating master clock is a hand-picked hydrogen maser, with outstanding short-term and intermediate-term stability. On the output of the free running maser is a state-of-the-art synthesiser, which is steered smoothly to a best estimate of UTC. This provides a real-time physical clock with excellent stability, and which can be seen, after the fact, to almost always be within 100 ns of UTC. See Chapter 6 for the details of the computation of UTC and TAI.

8.2.4.1 Models versus reality

Over the years, models of clock behaviour have been developed. These are very useful for simulation, for evaluating algorithms and for developing error detecting procedures, for example. The stochastic models may be stationary or non-stationary (flicker-noise FM for example), but this does not mean we can say the actual clock data are non-stationary. In reality, the data is only a composite of internal clock processes coupled to external perturbations and environmental conditions. Non-stationarity is a property of models, not clocks. The goal should be to obtain parsimonious models which, as much as possible, are explained by the underlying physics of the clocks and their interaction with the real world.

8.2.4.2 Data formats

Because clocks are often compared internationally, the BIPM has set up a standard procedure for providing input data for UTC. In addition, to get the most time transfer accuracy from GPS used in the common-view mode, a new standard has just been published in Metrologia. Those timing centres wishing to contribute need to contact the BIPM and comply with these standards. Some confusion has arisen in the past, for example in using GPS data, whether the data were taken on a 1 ns basis or 0.1 ns basis. The new format is 0.1 ns.

8.2.4.3 Retrieving and storing data

In F/T metrology, data are often taken continuously. If the data rates are high and the data words long, files can become massive – even with current large data file storage capability. The architecture of a system needs to be carefully thought out in terms of measurement noise, real-time timing needs, computation speed and memory size, F/T error detection, robustness and reliability, and accuracy and stability goals. The best time-scale system requires access to past data. For example, the measurement noise may be so high that it is necessary and sufficient for much of the time scale work needed to take data only once a day. The interplay between current data and the past history of contributing clocks is an important part of F/T metrology. We suggest contact with those who have had a lot of experience in this area before designing a system.

8.2.4.4 Installation concerns

The following concerns should be considered when installing an F/T system. Minimise ground-loop conditions when providing electrical power to various parts of the system, which is especially critical between a low-noise measurement system and the clocks. Magnetic fields should be minimised – AC (*e. g.* 60 Hz or 50 Hz) and changes in DC levels. Some magnetic shielding material used prudently can help. Having a temperature chamber for the clocks is a pattern followed in many of the best timing centres, ± 0.1° C control is quite useful. Areas of vibration should be avoided. A power conditioner for power-line control is helpful, and an uninterruptable power supply (UPS) is essential. VSWR issues for all critical RF cables is very important. The cables between the clocks and the output amplifiers should have isolation levels of more than 110 dB. Some of the most recent clocks have excellent environmental insensitivity.

8.2.4.5 Care and replacement

Monitoring of critical clock parameters is important in a long-term operating system. This helps in assessing procedures for repairs and in the decision making process for parts replacements. Most clocks have very finite lives and an operating budget should include a reasonable replacement schedule.

8.3 Conclusion

Frequency and time devices offer good opportunities for a variety of applications. Confidence in the measurements results is largely increased if the user applies a few common sense rules: redundancy, care in procedures and measurements, estimates of all the uncertainties and advice from the metrological centres.

References

AUDOIN C. and DIENER W.A. [1992] Frequency, Phase and Amplitude Changes of the Hydrogen Maser Oscillation. Proceedings of IEEE Freq. Control Symp., p. 86.

BIPM [1993] Annual Report of the BIPM Time Section.

CCIR [1990] Report 898-2: Operational Experience with Reference Clocks in Time Systems. Reports of the CCIR, Annex to Volume VII, Standard Frequencies and Time Signals, Dusseldorf, pp. 179-188.

FREON G. [1990] Preventive Maintenance and Stability of Commercial Caesium Clocks. Proceedings of 4^{th} EFTF, p. 549.

LEPEK A. and WALLS F.L. [1993] Cross Correlation Analysis Improves Time Domain Measurements. Proceedings of Freq. Control Symp., pp 313-320.

LEWANDOWSKI W. and THOMAS C. [1991] GPS Time Transfer. Proceedings of IEEE, Special Issue on Time, Vol. 79, p. 991.

LEWANDOWSKI W. and TOURDE R. [1990] Sensitivity to the External Temperature of some GPS Time Receivers, Proceedings of 22nd PTTI, p. 307.

NIST [1990] Characterization of clocks and oscillators. NIST Technical Note 1337, March.

PERCIVAL D.B. and WINKLER G.M.R. [1975] Timekeeping and the Reliability Problem. Proceedings of 29th Freq. Control Symp.

STEIN S.R. [1985] Frequency and Time, Their Measurement and Characterization. Precision Frequency Control, Vol. 2, p. 191, edited by E.A. Gerber and A. Ballato, Academic Press, New-York.

Chapter 9

Future prospects

Contents

9.1 Introduction

Descriptions, principles of operation and present performance for the well known existing atomic frequency standards: cæsium beam, rubidium gas cell, and hydrogen maser, have already been presented in Chapters 1 and 2 of this handbook so the descriptions given here will be very brief. This chapter covers what might be expected from these standards in the future as well as some of the newer standards under development that are likely to become available in the future. The new devices will be described and estimates of their performance will be given. These new standards include trapped ion devices, caesium fountain standards, oscillators stabilised to GPS (Global Positioning System) and devices in the optical frequency range. In addition, precision quartz oscillators will be discussed since they are essential elements of the atomic, ionic and GPS stabilised standards. Cryogenic dielectric resonator stabilised oscillators will also be described. They have extremely good short-term stability and could become important elements in the highest performance standards of the future.

9.2 General overview

There has been much progress recently in many areas of the frequency standard arena and the future prospects are quite promising, both in laboratory and commercial standards. Telecommunication and data communications are becoming very important and are driving many aspects of the commercial frequency standard business. Crystal oscillators and gas cell standards stabilised to GPS will be widely used in moderately demanding applications. In the laboratory area, the quest for higher accuracy and stability continues.

The main objectives for future commercial standards include lower cost, smaller size, higher performance and improved reliability. In particular, lower cost and smaller size are extremely important especially in high usage applications and therefore continued progress in micro-miniaturisation and increased integration of electronics is crucial. Better performance in both stability and accuracy will always be needed and considerable progress is being made here.

The efforts in laboratory standards are directed toward reducing or eliminating systematic effects thus improving the potential accuracy and long-term stability. Improved short-term stability is being pursued since it is needed to verify accuracy and stability in reasonably short measurement times, in addition to those needs for good spectral purity.

9.3 Gas cell devices

As mentioned in Chapters 1 and 2, gas cell frequency standards work by passing a beam of pumping light through a gas cell containing vapour of the atom being used (typically rubidium or cæsium) with, usually, a buffer gas in an excited microwave cavity. The system is designed so that the intensity of the pumping light transmitted through the cell is a minimum when the microwave excitation frequency is at the atomic resonance.

The quantity of present rubidium standards sold is considerably larger than any of the other atomic frequency standards primarily due to cost and size. They are becoming an important element in telecommunications. Their performance is typically between quartz and cæsium beam standards. Properly designed units can have considerably better shock and vibration performance than quartz. The same is true regarding shifts due to change in orientation in the earth's gravitational (g)-field. Cost and size of these units are of prime importance and these are continually being reduced.

Most of the present rubidium units are optically pumped with an RF excited lamp. Typical performance is shown in Table 9.1. A future development is to replace the RF excited lamp with a suitable solid-state laser. Using laser pumping can lead to smaller size and perhaps to the performance shown in Table 9.1.

TABLE 9.1

Performance of compact Rb standards

Parameter	Performance (compact Rb with lamp)	Performance (compact Rb with laser)
Ageing magnitude (10^{-11} /month)	1 to 2	1 to 2
g Sensitivity magnitude (10^{-11}/g)	<1	<0.2
Flicker floor (10^{-13})	3 to 5	1
Temperature sensitivity magnitude (10^{-13}/ C)	<6, (but non-linear)	<1
Spectral purity, dBc in 1 Hz bandwidth at: 1 Hz 10 kHz	 –80 –145	 –80 –145
Short-term stability: 10^{-12} (1 s avg. time) 10^{-13} (100 s avg. time)	 3 3	 <1 <1
Volume in^3 cm^3	 16 260	 6 100

Using a laser for pumping a rubidium cell and designing the system for optimum performance rather than small size can give short term stability perhaps as good as $2 \times 10^{-14} \times \tau^{-1/2}$, where τ is the averaging time. A very low flicker floor may also be achieved. This type of device is a good candidate for a flywheel source for advanced very-high-stability atomic standards and consequently of potential importance.

Work is presently being done on a laser pumped cæsium gas cell device. This standard could be somewhat smaller than RF lamp pumped rubidium devices mainly because of the shorter wavelength of the cæsium line, 3.26 cm versus 4.39 cm for rubidium, and the small size of the laser compared to the RF excited lamp. The performance should be comparable to rubidium but the ageing may be poorer due to the relatively high surface-area-to-volume-ratio in a small cell. A highly integrated set of electronics along with the small physics package could reduce the volume to 10 cm^3 or less and perhaps lead to lower cost if the manufacturing volume is large. These standards could have wide use in the future.

The already large market for commercial gas cell devices has the potential to grow even larger if cost and size can be brought down. Laser pumping is important for both size reduction and performance. It could even reduce the cost compared to RF excited pumping lamps if the laser cost is reasonable. Laser availability and price are crucial and certainly depend on having large unit volume, so how this overall situation develops in the future is uncertain. There is some indication that at least one laser manufacturer is somewhat interested in the laser pumped frequency standard business. Laser pumping can produce dramatic improvements in short-term stability over conventional gas-cell Rb standards.

9.4 Cæsium beam standards

As described in Chapters 1 and 2, cæsium beam standards work by passing a beam of state-selected cæsium atoms through an excited microwave cavity. On leaving the cavity, the atoms are subjected to further state selection to choose those that have made a microwave transition and eventually obtain a signal that is maximum when the microwave excitation frequency equals the resonance frequency of the atoms. Some form of frequency or phase modulation of the microwave excitation is always used to allow the precise determination of the line centre.

Cæsium beam frequency standards are important where high accuracy, reproducibility and negligible drift are needed. The present highest performance commercial unit has an accuracy magnitude better than 1×10^{-12}, a drift magnitude much less than 1×10^{-15} per day, flicker floor less than 1×10^{-14}, short term stability better than $8 \times 10^{-12} \times \tau^{-1/2}$, and a temperature coefficient magnitude less than 1×10^{-15} per degree C. The high performance variety of cæsium beam standards is moderately expensive.

Optical pumping of cæsium beam devices using lasers to achieve state selection and atom detection is going on in a number of laboratories at the present time. The new laser pumped standard at NIST (NIST-7) is now operational and giving outstanding performance. Several other laboratories are working on optically pumped cæsium beam tubes at the present time including some work on small tubes for commercial application. Laser pumping to applied commercial standards will improve their accuracy by perhaps a factor of 3 to 5 and short-term stability by more than 10. Accuracy improvement is due to several things. Rabi and Ramsey pulling can be significantly reduced and the C field homogeneity is better. These are due to the lack of deflection magnets in the optically pumped tube and the improvement in symmetry of the microwave transitions close to the main transition if the right type of pumping is used. In addition, better correction can be made for the frequency shifts due to cavity phase shift and relativity (the second order Doppler shift). Short-term stability improvement comes about because of the much better utilisation of the cæsium in the beam. Improved short-term stability is particularly important since that is the weakest performance area in present commercial cæsium beam standards. Improving the short-term stability by 10 reduces the time to make a measurement to a given precision by 100 – it would take an ensemble of 100 unimproved standards to get to the same precision in the same time! The availability of high reliability lasers is again crucial here.

The market for lower cost, and consequently lower performance, cæsium beam standards may grow due to continual increases in timing and synchronisation requirements as communication rates go up. However, GPS stabilised quartz oscillators, as described later in the chapter, could take over a significant portion of this market.

9.5 Hydrogen masers

Active hydrogen masers described in Chapters 1 and 2 utilise the stimulated emission of hydrogen atoms in a cavity to produce an actual oscillation at the hydrogen hyperfine frequency, 1 420 MHz, in contrast to the passive standards we have discussed so far. They provide the best short-term stability presently available from an atomic standard in the microwave range. Typical performance, using mod $\sigma_y(\tau)$ as the stability measure, is about $1 \times 10^{-13} \times \tau^{-3/2}$ for times shorter than about 20 seconds and $2.2 \times 10^{-14} \times \tau^{-1/2}$ till the flicker floor or drift is reached. The best stability reached is typically somewhat better than 1×10^{-15}. Active hydrogen masers are the standards of choice when extremely good short term stability is required such as in Very Long Baseline Interferometry, and in other radio astronometric applications.

There is frequency pulling of the maser due to cavity mis-tuning leading to frequency drift as the cavity drifts with time. There are, however, several techniques for auto-tuning the cavity that effectively remove this source of drift. Magnitude of drift rates of units without auto-tuning of the cavity normally range from 1×10^{-15} per day to 1×10^{-14} per day.

There is also a frequency shift due to collisions of the hydrogen atoms with the storage bulb walls that amounts to a magnitude of about 2×10^{-11} for typical masers. Due to lack of precise knowledge of this wall shift, the absolute accuracy magnitude of the hydrogen maser is presently limited to about 1×10^{-12}.

Active hydrogen masers are relatively expensive and the market for them is not large at this time.

Passive hydrogen masers are similar to the gas cell devices and the cæsium beam standards already discussed. Their short term stability is considerably poorer than active hydrogen masers but somewhat better than present high performance commercial cæsium standards. Passive masers have the same wall shift uncertainty as the active maser but cavity pulling is much smaller in common with the other passive standards. Work has been done in the U.S. for some time on passive masers but no U.S. commercial units are on the market. However, commercial units are presently being soldsale by a Russian firm. The market is not large for these standards either.

Work has been going on in several places on cold (cryogenic) active hydrogen masers. These are expected to have extremely good short term stability, achieving perhaps better than 1×10^{-18} at around 1 000 seconds, and very good stability with respect to ambient temperature changes. The required refrigeration is fairly complex and therefore these masers would be fairly expensive. They could also be an excellent flywheel oscillator for some of the advanced passive standards.

9.6 Trapped ion standards

Trapped ion standards are passive devices that use an RF quadrupole structure (Paul trap) to trap quantities of one to many (about 1×10^7) ions. Two types of trap are used, one with ring and cap electrodes and the other with four symmetrically placed rod electrodes and end electrodes. A single ion with the proper charge to mass ratio in the combined RF and DC fields of either of these traps experiences an average force directed towards the geometric centre of the structure. If the ion is viscously damped it will come almost to rest there. If many ions are trapped, the combined fields of the quadrupole and the space charge of the ions lead to the formation of an ion cloud if there is some cooling or viscous damping. The ring and cap structure can sustain a spherical cloud while the rod structure sustains an elongated cloud. In any case, ions so trapped can be interrogated for very long periods of time thus leading to very narrow resonance lines.

One technique used for viscous cooling is to introduce helium into the trap at low pressure (1×10^{-6} torr). The heavy ions vibrating in the RF trap field can then lose thermal energy by colliding with the light helium atoms. Laser cooling is another technique that is being explored with good success.

In the mercury 199 version of the trapped ion standard, optical pumping using an RF excited lamp or a laser source is used for state preparation and observation of the microwave resonance at 40 GHz. Line widths well below 0.1 Hz have been achieved. The biggest systematic frequency offset when many ions are in the trap is due to the relativistic velocity effect (second order Doppler shift) caused by the induced motion of the ions in the RF trapping field. For a spherical cloud containing about 2×10^6 ions, the shift is about -2×10^{-12}. This is greatly reduced by using the elongated clouds sustained in the rod type trap or virtually eliminated by using a single ion or a line of single ions. If helium cooling is used, there is also a collision induced frequency shift of about 1×10^{-13}, the actual size of which can be determined by extrapolating experimentally to zero helium pressure.

Several trapped ion standards have been built using mercury 199 ions with an RF excited lamp for optical pumping. At least two groups are actively working in the field at present. Excellent short term stability, about $1 \times 10^{-13} \times \tau^{-1/2}$, has been demonstrated with an RF lamp pumped, elongated mercury ion cloud trapped in a two-dimensional rod type quadrupole trap.

Background gas molecules such as heavy hydrocarbons can cause fairly large frequency shifts so cryogenic operation may be required for the ultimate in stability. Calculations indicate that a single trapped ion using laser cooling and cryogenic pumping to get rid of background gas could have offsets from the free ion resonance frequency as low as 1×10^{-17}.

Trapped ion standards without either cryogenic cooling or laser optical pumping could be only somewhat more expensive than a high quality cæsium standard. Adding laser pumping for mercury ions is very expensive and bulky with the present state of the art.

No units are available commercially at this time. A version using an RF lamp pumped elongated mercury 199 cloud and low pressure helium gas cooling could be made commercially and would have performance superior to high quality cæsium beam standards at not too great a cost.

9.7 Cæsium fountain

The cæsium fountain is a passive standard that uses laser cooling and manipulation to toss a ball of extremely cold (microKelvin level), state-selected cæsium atoms upwards through an excited microwave cavity. They then fall back through the same cavity under the influence of gravity and their state is evaluated to see if the microwave transition has occurred. The atoms see a double pulse of microwaves which is the usual Ramsey excitation. Transit times as long as 1 second between the two interactions with the cavity have been achieved, leading to a line-width as narrow as 0.5 Hz with good signal-to-noise ratio. Several groups are presently working on these devices and progress is rapid.

Accuracy is expected to be 1×1^{-15} or better. The main limitation is a density dependent frequency shift due to spin exchange collisions. This shift can be quite large at the low temperatures and moderate densities used. Extrapolation to zero density can be done by a series of measurements and the accuracy value given above assumes this procedure has been carried out. Short term frequency stability of $3 \times 10^{-14} \times \tau^{-1/2}$ is obtained with the signal-to-noise ratio already achieved. Acceleration and apparatus orientation effects will be large with such slow atoms. Thus, the cæsium fountain standard will only work in an environmentally benign location. No commercial units are presently available.

9.8 Quartz oscillators

Performance of present high quality quartz oscillators is limited by the resonator and environmental control of the resonator and associated critical circuitry. BVA crystals, made from an all quartz structure with electrodes spaced from the crystal surface, are the best resonators available today from the standpoint of frequency stability and drift. Unfortunately, the manufacturing costs are higher than conventional resonators with electrodes formed directly on the crystal.

Precision quartz oscillators are important not only for stand-alone applications but also because they are the essential output elements of most of the atomic standards. In passive standards, they provide the basis for the microwave interrogation signal. They are essential for the active hydrogen maser since the actual hydrogen oscillation power level is very low. In either case, the frequency behaviour of the standard for long times is that of the atomic resonance or oscillation while at short times the behaviour is governed entirely by the quartz flywheel. The time constant at which the crossover occurs depends on the noises of the atomic reference and the quartz oscillator and is chosen to optimise overall performance. The better the oscillator, the better the overall performance will be.

In passive standards, noise at the second harmonic of the frequency used for the intentional frequency or phase modulation to find the line centre causes additional noise that is indistinguishable from the noise of the atomic resonator. This is an important limitation in high performance standards. One technique to reduce this effect has recently been developed at NIST. Two notch filters, one on each side of the oscillator frequency and separated from it by twice the intentional modulation frequency, serve to attenuate the noise modulation very effectively. The effect of the spurious noise can be further reduced by an appropriate signal processing of the atomic resonator response, as shown by a group in France.

Performances of several oscillators are shown in Table 9.2. The ultra-precision 5 MHz BVA oscillator has the best overall performance. The 10 MHz BVA is next. The third oscillator uses a 10 MHz conventional SC cut resonator. The magnitude of the ageing rate of the best oscillators is occasionally as low as 2×10^{-12} per day but more typical values are 10 to 50×10^{-12} per day. Occasionally flicker floor levels as low as 4×10^{-14} are seen. These results are mainly determined by the resonator as long as the circuitry is well designed.

Complete oscillator temperature sensitivity magnitudes range from 0.2 to 40×10^{-12} per degree C. Clearly the temperature results are quite variable depending on the design and construction of the oscillators and the necessary oven containing the crystal and some of the circuitry.

TABLE 9.2

Performance of quartz oscillators

Parameter	5 MHz Ultra-Precision	10 MHz Ultra-Precision	10 MHz Precision	Future Ultra-Precision
Ageing magnitude (10^{-12}/day)	2 to 50	5 to 50	<100	1
g Sensitivity magnitude (10^{-10}/g)	1	0.5	<10	<1
Flicker Floor (10^{-13})	0.4 to 2	1 to 3	10	0.1 to 0.5
Temperature Sensitivity magnitude (/ °C)	2 to 5	<50	<400	<1
Spectral Purity dB in 1 Hz Bandwidth at: 1 Hz 100 kHz	 –130 –180	 –120 –155	 –105 –162	 –145 –165

The fact that very good ageing and flicker performance can be obtained occasionally indicates that considerable improvement in resonators is possible but the processing and/or material is presently not well enough controlled.

Advances needed in quartz oscillators include: better material for the resonators perhaps accompanied by higher intrinsic *Q*; improved understanding and technology of the quartz-electrode interface; better oven designs and hermetic sealing to reduce environmental effects. Ultimately, the electronics will need improving. The multiple series resonator approach recently patented by Westinghouse can give better overall performance and may be used in the most critical applications.

Table 9.2 also shows what might be expected in future ultra-precision oscillators.

9.9 Oscillator stabilised to GPS

If an oscillator is stable enough, it can be used in a phase-lock loop with a GPS receiver and can average out the frequency variations, i.e. selective availability, intentionally put on the GPS signals. If the oscillator is locked to the average GPS frequency with appropriate filtering algorithms, excellent long-term stability with respect to GPS can be achieved. Any high quality oscillator or standard can be used: quartz, Rb cell, Cs beam, hydrogen, etc.

RMS time uncertainty with respect to GPS of about 2 ns can be achieved with high quality cæsium and a good receiver. The uncertainty expected with a high quality quartz oscillator is perhaps 20 ns. The frequency stability of GPS itself is a limiting factor. Present measurements indicate that GPS has an Allan deviation of about 2×10^{-14} for one day averaging time. The time stability of GPS is also important. Present regulations allow an RMS offset of 340 ns but measurements give values about one-tenth this amount.

This technique is relatively low-cost when a quartz oscillator or gas cell standard is used. It is just becoming available commercially and could displace higher priced atomic standards in a number of applications. A cæsium standard or ensemble of cæsium standards locked to GPS with an appropriately long time constant would have exceptionally good performance.

9.10 Oscillator stabilised with cooled sapphire resonator

Cooled sapphire has extremely low dielectric losses at microwave frequencies and consequently a sapphire microwave dielectric resonator can have very high Q. Values in excess of 2×10^7 at X-band have been measured. Several groups are working on oscillators stabilised by such resonators and are achieving excellent results in the X-band region with temperatures of 77 Kelvin and below.

The spectral purity achieved at X-band is about –50 dBc referred to a 1 Hz bandwidth at 1 Hz offset from the carrier. The noise floor is presently about –162 dBc and is reached at about 3 kHz. These, particularly the noise floor, are much better than can be achieved with a high quality 10 MHz quartz oscillator multiplied in frequency to X-band. This stability at X-band has been exceeded only by an oscillator stabilised by a superconducting niobium cavity at about 2 degrees Kelvin. A source with this kind of spectral purity is needed for very high performance Doppler radars. It would also be useful as a high performance flywheel for advanced microwave and optical atomic standards. Very careful design and construction is necessary to avoid deleterious modulation by acoustic noise and vibration. Only one unit is available commercially at this time.

9.11 Optical frequency standards

Optical atomic frequency standards have the potential for very high accuracy and stability. Stabilised helium-neon lasers at 633 nm using the Lamb dip and other techniques have been commercial items for more than 25 years. Their reproducibility and accuracy is around 1×10^{-7}. Much better performance was obtained with lasers stabilised to iodine absorption lines. Methane stabilised lasers in the near infrared appeared with reproducibility of a few parts in 10^{12}. A great deal of work on methane was done in the U.S. at what was then NBS (now NIST) and also in the USSR (now Russia). A number of portable methane standards were built and used in the USSR.

Very good candidates exist among atoms and ions with very narrow optical spectral lines.

The mercury ion is a good example. Accuracy of standards built using these lines is predicted to be considerably better than 1×10^{-15}.

While these standards are excellent in the optical range, it is still very difficult, complicated, and expensive to connect their frequency with no loss in accuracy to the RF/microwave region. Work is being done in this interesting and challenging area in several locations and there are some promising ideas that may ultimately lead to a practical solution to the connection problem.

9.12 Summary

Continued improvement will be made in performance, size, cost, and reliability in many of the existing standards. Inexpensive, small gas cell devices will have a large market. Telecommunications and data communications are rapidly growing fields and their synchronisation requirements are getting tighter. These are strong driving forces for the frequency standard market. GPS stabilised oscillators will be widely used.

In the higher performance area a number of things are happening. Optically pumped commercial cæsium beam standards will offer considerable improvement in stability and accuracy. Trapped mercury ion devices have very good potential. Good flywheel oscillators will always be needed and their performance requirements are severe for the advanced atomic standards coming along. Very high performance standards such as the cæsium fountain and single ion devices are in active development and may be commercialised in the future. Optical standards have great promise but the problem of easy, accurate connection to the RF/microwave frequency range needs to be solved. Very high performance standards will always be needed for special applications and scientific work but their market is not large.

Chapter 10

Conclusions

Contents

10.1 General observations

The subjects of the Chapters in this Handbook are interdependent. Describing one topic without calling upon information from another is difficult. While the author of each chapter does not often refer explicitly to other chapters, the reader will no doubt find a need to do so. A key conclusion then is that the pursuit of work in one topic will probably require development of understanding of others. Consider the following examples of interdependence.

Advanced clocks and oscillators cannot be readily understood without an understanding of earlier clocks and oscillators, and work on either current or advanced devices surely requires an understanding of both the statistical basis for characterisation and methods of measurement.

To appreciate time and frequency applications, an understanding of several topics might be needed. These include the characteristics of clocks and oscillators, the construction of reliable time scales, the dissemination of time and frequency signals, the characterisation of signals transferred through such dissemination systems, and practical problems associated with these activities. For example, the performance of received signals often depends on the characteristics of the local broadcast source, the time scale to which that source is steered, and the dissemination process.

10.2 Clocks and oscillators

The range of clocks and oscillators available for applications is excellent. Commercially available clocks and oscillators include, for example; quartz, rubidium, cæsium, and hydrogen devices. This is the approximate ascending order of cost of these timing devices. Their basic principles of operation (outlined in Chapter 1) and their performances (outlined in Chapter 2) vary considerably. Specific application requirements (frequency stability, drift, sensitivity to environment, size, cost, weight, power requirements, etc.) might suggest selection of any one of these devices. Beyond these current commercial solutions to application requirements lie research advances (outlined in Chapter 9) that suggest future availability of yet more advanced devices. The paragraphs below summarise some key conclusions in this area.

Quartz oscillators form the very foundation of time and frequency technology. Passive atomic standards are critically dependent on quartz oscillators. In fact, the attack time of the servo systems in these standards is long. This means that the short term stability of passive atomic standards is essentially that of the quartz oscillator used to control the microwave oscillator that probes the atomic resonance. The number of quartz oscillators used in even very high-technology applications is enormous. They clearly dominate atomic standards in numbers by several orders of magnitude. While their long-term stability and drift are inferior to those of atomic standards, quartz devices remain desirable solutions because of their low cost, low weight, small size, high reliability, and low power requirements.

Rubidium frequency standards occupy the next level up in performance. Their stability is usually better than that of quartz oscillators for times longer than 10^4 s, but is not nearly as good as that of beam standards. They are generally less sensitive to environmental changes than quartz. While the packaged rubidium standard is typically larger than a packaged quartz oscillator, it is much smaller than currently available cæsium beam standards. Rubidium frequency standards are finding ever more applications, particularly in telecommunications systems, and have performed well as on-board clocks on the GPS satellites.

Cæsium beam frequency standards occupy a unique niche, because they rely on the resonance that serves as the definition of the second. Consequently, cæsium standards are the only choice for national primary frequency standards. The primary frequency standard differs from the typical field (commercial) standard in that systematic errors are carefully evaluated to arrive at the best possible realisation of the second. In practice, many systematic effects are under good control in commercial cæsium beam standards, so they can provide accuracy, but at a level below that of the primary standards. Because systematic effects are under better control, cæsium beam standards typically exhibit less drift than other standards, and the user can have more confidence that the output frequency corresponds closely to that stated by the manufacturer. Thus, cæsium beam standards are usually the first choice for systems where accurate frequency must be autonomously kept for long periods. Cæsium beam standards play a major role in the ground-control stations of GPS and are also carried (along with rubidium standards) by the GPS satellites. At present (1995), most of the clocks providing time and frequency from the GPS satellites are cæsium.

While priced much higher than other frequency standards, hydrogen masers deliver far superior stability in the short and intermediate term. They yield only to cæsium in the long term, and even then not always. If the resonance of the microwave cavity is servo controlled, a hydrogen maser can be more stable than a commercial cæsium standard for times of the order of one year. The long-term-stability problem is a result of the difficulty in controlling systematic effects. One of the most troubling of these is the shift associated with collisions of the hydrogen atoms with the walls of the storage bulb. Another is the pulling of the hydrogen-frequency resonant cavity, causing an error in the output frequency. The short-term stability of hydrogen masers has been extremely useful in applications such as very long-baseline interferometry (VLBI) where the time tagging of closely spaced observations is especially critical. Masers are also useful for characterising other high-performance frequency standards, since the time needed to make a given measurement becomes smaller as noise in the reference device decreases. The typical hydrogen maser is active, that is, it oscillates spontaneously, but passive hydrogen masers have also been built. Passive masers are intermediate in performance between active hydrogen masers and cæsium standards.

Research is under way on all these standards. Quartz devices have been improving gradually with time, and gradual improvement in the future is likely to follow improvement in understanding of materials and noise mechanisms, and development of clever methods for controlling environmental sensitivities. On the other hand, research in atomic physics has provided new tools that promise improvements of several orders of magnitude in performance of atomic standards. Lasers can now be used to control the atomic states and motions of atoms, dramatically reducing limitations related to atomic motion (Doppler shifts) and observation time. The ion frequency standard and the cæsium fountain top the list of advanced passive standards. Rubidium standards will likely benefit from laser pumping techniques that could replace discharge-lamp pumping, since performance limitations are ascribed to the traditional pumping method. Hydrogen masers should benefit from research on cryogenic wall coatings that promise substantial improvement in their already impressive short-term stability. Finally, mention should be made of research on the cooled sapphire resonator and the superconducting microwave resonator, both of which can exhibit exceedingly high Q, consequently leading to short-term stability well beyond that of the hydrogen maser.

10.3 Measurement methods and characterisation

Measurement methods in this field are dictated by the nature of the noise encountered in clocks, oscillators, and signal transfer systems. The realisation that lower-frequency noise is not usually white, but varies as the reciprocal of some higher power of the frequency, demands non-standard statistical treatment of the noise. Since timekeeping is a long-term activity requiring consideration of very long-term (low-frequency) behaviour, the statistical treatment of these non-white noise processes is particularly important. The random noise is clearly important, but systematic effects must also be understood, characterised, and controlled. Over the last several decades these subjects (well described in Chapters 3, 4, and 5) have evolved substantially.

Both time-domain and frequency-domain characterisation methods are important. Time-domain measures (described in Chapter 3) are especially useful for characterising long-term processes and frequency-domain measures are more useful for characterising shorter-term (higher-frequency) behaviour. The two-sample (Allan) variance and variations of it have satisfactorily replaced the standard variance, which cannot be used because it diverges for some non-white noise. Two-sample variances depend on the averaging time, so this characterisation approach involves a graph, a $\sigma - \tau$ plot, rather than just a one-number variance. The variance measures lend themselves well to the technology, since repeated measurements with a counter can be used to acquire data. Averaging time τ is then adjusted in the software processing of the data. The τ dependence is related to the Fourier frequencies present, hence a $\sigma - \tau$ plot can give a quick indication of the frequency components present in the data.

Shorter-term noise is more typically described using a spectral purity measure such as power spectral density. Phase-modulation (PM) noise is typically of greatest concern, but amplitude-modulation (AM) noise should not be completely overlooked. Chapters 3 and 4 provide descriptions of both the characterisation concepts and the methods for making the physical measurements. Direct measurements using a spectrum analyser can be useful, but the very high performances of many systems require higher resolution, which is usually achieved using a heterodyne method to bring the noise down from the higher frequency to base band.

Clocks and oscillators are subject to systematic effects induced by variations in such standard environmental parameters as temperature, humidity, barometric pressure, and magnetic field (see Chapter 5). Acceleration, vibration, shock, and ageing are also of concern. The physical origins of the responses of the various types of standards to changes in these environmental parameters are obviously different. An understanding of the relative magnitudes of these effects in the various devices can be of help in selecting a clock/oscillator for a particular application or in taking necessary precautions to reduce the effects of variations in the environment. It is also useful to understand the origins of these systematic effects. This requires study of their physical principles (Chapters 1 and 2).

Characterisation of clocks and oscillators is multifaceted, and errors are often made. A detailed understanding of the statistical measures and the measurement methods is essential to good metrology. Errors can result from mistakes in zero-crossing detection, truncation of data, improper accounting for bandwidth effects, and improper consideration of counter dead time. Measurements can be mathematically converted between the time and frequency domains, but considerable caution must be exercised in doing so, especially when going from the time to the frequency domain. Misunderstanding of characterisation concepts has led to so many problems that several organisations have developed standard nomenclature for terms and standard methodologies for performing measurements.

10.4 Time scales, coordination, and dissemination

For many applications the usefulness of the frequency output of an oscillator or the time output of a clock is not realised unless that signal is reliably maintained, compared with other oscillators/clocks and distributed to other locations (network-type applications). Many applications require continuous availability of time and/or frequency traceable to a central source. The questions that naturally arise involve methods for reliably maintaining the sources and optimum means for distributing signals from those sources.

In recent years, the ensembling of clocks with a proper algorithm to form a time scale has emerged as an important concept (Chapter 6). A properly implemented time scale can provide higher reliability and performance beyond that of any of the clocks contributing to the scale, and the means for assessing the performance of component clocks and even the time scale itself. Ensemble time scales rely on many earlier developments including methods for clock characterisation, low-noise methods for reading the outputs of clocks, and algorithms for optimally combining the outputs of an ensemble of clocks.

To maintain an autonomous time scale with an output related to UTC or any other time scale, co-ordinating the time-scale with the reference through time transfer is critical. Satellite methods have dramatically improved Coordination to the point where the best standards and time scales can be compared at their full accuracy, although this might require a number of days of averaging for the highest performance systems. Thus, the highest available accuracy can now be transferred anywhere. If an application requires continuous local availability of time, maintaining a high-stability clock system (preferably an ensemble) is best. The system can be steered with a long time constant to an external source. If the external signal is lost for a period the local system can adequately bridge the gap.

Most requirements for time and frequency signals involve accuracies well below those of the major reference time scales and simpler, more cost-effective methods of delivery of these signals (as described in Chapter 2) can be used. Several different signal delivery systems have been used over the years. These include telephone, HF radio, LF radio, and satellite signals. These are typically operated by national standards laboratories. In addition, other stable signals, not usually considered as standards, are often used as frequency references. A standard's laboratory can monitor such signals and provide correction data in the form of offsets from their own standard. The user applies these corrections to the received signal to achieve an accuracy that is traceable to the reference standard. Television and LORAN-C broadcasts have often been used in this way. Such methods allow a national laboratory to provide excellent service without a large investment in broadcast equipment.

The need for time stamping of business transactions and technical data by distributed computer systems has been growing rapidly. Current telephone and network time services are meeting some of these needs, but it is likely that timing signals will eventually be provided for such applications by telephone companies through fibre and cable systems.

10.5 Realities

Applications typically drive the development of a technology, and this is certainly the case (as discussed in Chapter 7) for time and frequency technology. Applications range from the very demanding scientific experiments to clocks and oscillators for a variety of simpler consumer products. Clocks and oscillators control the activities of many systems, that is, they regulate the rate at which events take place. As our societies have become more dependent on electronics, timing of activities has become progressively more exacting.

Frequency is our most accurately realised physical quantity. Because frequency measurements of modest accuracy are easy to make, other quantities are often transduced (converted) to frequency. For example, length is now defined as the distance travelled by light in a fixed time. Also, the Josephson volt standard converts frequency to voltage. These are examples of conversions of major measurement units, but many other types of transducers convert other quantities (for example, pressure and vibration) to frequencies. Thus, the value of frequency standards extends well outside the traditional areas of timing.

The close relationship between length and time has been exploited in navigation systems; the Global Positioning System (GPS) is particularly noteworthy. An accurate time interval translates directly to an accurate distance. In fact, the development of atomic frequency standards has been critical to the development of these navigation systems. The impact of these developments has been large. GPS, for example, is used not only for navigation by ships and aircraft, but for construction surveying, and in geophysical research. The ability to determine position accurately will no doubt be applied to many other activities generating a wide range of consumer products for positioning.

Telecommunications is another key area of application. Timing requirements are escalating as systems move to ever higher data rates. With the explosion in information transfer and processing, time and frequency technology is playing a larger role in telecommunications. The capabilities of atomic frequency standards have long been required in this industry, and wide distribution of high-level timing is critical to the successful operation of newer systems.

While the list of applications goes on, it might be well to conclude here by noting that this technology is also important to scientific research. Many fundamental physical theories yield predictions of clock behaviour, and clocks thus play a major role in tests of these theories. In fact, scientific applications have often served as the driving force for the development of better clocks.

This Chapter is concluded with a few words of caution. This is a complex technology involving a range of disciplines. These include, for example, atomic physics, materials science, electronics, measurement science, statistical analysis and satellite and terrestrial radiocommunications. Stability, reliability, and accuracy have come to have special meanings and importance in the field. These must all be solidly understood. Practitioners of the technology have accumulated experience suggesting that certain types of problems occur repeatedly. Some of these problems are discussed in Chapter 8. If you are about to embark on a major project in the field, seek the experience of others. Chapter 8 might provide a good starting point.

10504

Printed in Switzerland
Geneva, 2008
ISBN 92-61-06511-2