

Dennis D. McCarthy and
Kenneth P. Seidelmann

 WILEY-VCH

TIME

From Earth Rotation to Atomic Physics



Dennis D. McCarthy
P. Kenneth Seidelmann

**TIME – From Earth Rotation
to Atomic Physics**

Related Titles

F. Riehle

Frequency Standards

Basics and Applications

2004

ISBN: 978-3-527-40230-4

H. Mark (Ed.)

Encyclopedia of Space Science and Technology

2 Volume Set

2003

ISBN: 978-0-471-32408-9

V.G. Oklobdzija, V.M. Stojanovic, D.M. Markovic, N.M. Nedovic

Digital System Clocking

High-Performance and Low-Power Aspects

2003

ISBN: 978-0-471-27447-6

S. Bregni

Synchronization of Digital Telecommunications Networks

2002

ISBN: 978-0-471-61550-7

Dennis D. McCarthy and P. Kenneth Seidelmann

TIME – From Earth Rotation to Atomic Physics



**WILEY-
VCH**

WILEY-VCH Verlag GmbH & Co. KGaA

The Authors

Prof. Dennis D. McCarthy

U.S. Naval Observatory
Washington DC, USA
dennis_mccart57@hotmail.com

Prof. P. Kenneth Seidelmann

University of Virginia
Astronomy Department
Charlottesville, VA, USA
pks6n@virginia.edu

■ All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>

© 2009 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Cover picture Spiesz Design, Neu-Ulm
Typesetting SNP Best-set Typesetter Ltd., Hong Kong
Printing betz-druck GmbH, Darmstadt
Bookbinding Litges & Dopf Buchbinderei GmbH, Heppenheim

ISBN: 978-3-527-40780-4

*In honor of our wives, Diane McCarthy and Bobbie Seidelmann,
and our families.*

Contents

Preface *XV*

Acronyms *XVII*

1	Time: Pre-Twentieth Century	1
1.1	In the Beginning	1
1.2	Characterizing Time	1
1.3	Calendars	2
1.4	Astronomical Observations	3
1.5	Timekeeping	4
1.6	Time Epochs	5
1.7	Time Transfer	6
1.8	Rotation of the Earth	7
1.9	Beginning the Twentieth Century	7
	References	8
2	Solar Time	9
2.1	Apparent Solar Time	9
2.2	Mean Solar Time	9
2.3	Equation of Time	11
2.4	Sidereal Time	11
2.5	Washington Conference of 1884	12
2.6	Newcomb's Theory of the Sun	13
2.7	Universal Time	14
2.8	Coordinated Universal Time (UTC)	16
2.9	Greenwich Mean Time (GMT)	17
2.10	Tropical Year	18
2.11	Besselian Year	19
2.12	Reference System	19
2.13	Time Zones	20
2.14	Daylight Saving Time	20
	References	22

3	Ephemerides	23
3.1	Ephemerides and Time	23
3.2	Before Kepler and Newton	23
3.3	Kepler and Newton	26
3.4	Tables, General Theories, and Ephemerides	27
3.5	Lunar Theories	29
3.6	The Advent of Computers	32
3.7	Numerical Integrations	33
3.8	Observational Data	33
3.8.1	Radar Observations	34
3.8.2	Lunar Laser Ranging	34
3.8.3	Spacecraft Observations	34
3.9	Dynamical Reference Frame	34
3.10	Time Arguments	35
3.11	Astronomical Constants	36
3.12	Artificial Satellite Theories	36
3.13	Theory of Relativity	37
	References	37
4	Variable Earth Rotation	41
4.1	Pre 19th Century	41
4.2	Secular Variation	42
4.3	Irregular Variations in the Earth's Rotation	44
4.4	Early Explanations for the Variable Rotation	52
4.5	Current Understanding of the Earth's Variable Rotation	53
4.6	Consequences	55
	References	57
5	Earth Rotation and Polar Motion	61
5.1	Earth Orientation	61
5.1.1	Precession/Nutation	65
5.1.2	Polar Motion	66
5.1.3	UT1	68
5.2	Variations in the Earth's Orientation	69
5.2.1	Precession Nutation	69
5.2.2	Polar Motion	70
5.2.2.1	Secular Variation	70
5.2.2.2	Decadal Variations	70
5.2.2.3	Chandler and Annual Variations	71
5.2.2.4	Other Variations	71
5.3	Transforming Between Reference Frames	72
5.4	Determination of Earth Orientation	75
5.5	Earth Orientation Data	76
	References	77

6	Ephemeris Time	79
6.1	Need for a Uniform Time Scale	79
6.2	Danjon Proposal	80
6.3	Clemence's Proposal	81
6.4	Adoption and Definition	83
6.5	Observational Determination	84
6.6	The Ephemeris Second and Atomic Time	86
6.7	Historical ΔT	87
6.8	Problems with Ephemeris Time	90
6.9	Relativity	91
6.10	Dynamical Time Scales	92
	References	93
7	Relativity and Time	95
7.1	Newtonian Reference Systems	95
7.2	Special Relativity	95
7.3	Lorentz Transformations	97
7.4	Coordinate and Proper Time	98
7.5	Minkowski Diagrams	99
7.6	Time in Special Relativity	102
7.7	General Relativity	103
7.7.1	Metrics in General Relativity	103
7.7.2	The Equivalence Principle	104
7.8	IAU Resolutions	105
7.9	Time Scales	111
7.9.1	International Atomic Time	111
7.9.2	Dynamical Time Scales	111
7.10	Relativistic Effects in Time Transfer	111
	References	112
8	Dynamical and Coordinate Time Scales	113
8.1	Replacing Ephemeris Time	113
8.2	Terrestrial Dynamical Time (TDT) and Barycentric Dynamical Time (TDB)	114
8.3	Problems with TDT and TDB	116
8.4	New Reference System	117
8.5	New Time Scales	118
8.5.1	Coordinate Time	119
8.5.2	Terrestrial Time	119
8.5.3	Geocentric Coordinate Time	120
8.5.4	Barycentric Coordinate Time	120
8.5.5	Barycentric Ephemeris Time	122
8.5.6	TDB Redefined	122
8.6	ΔT and Ephemeris Time Revised	124
8.7	Relationships Among Coordinate Time Scales	125
	References	126

9	Clock Developments	129
9.1	Introduction	129
9.2	Keeping Time in Antiquity	129
9.2.1	Clepsydrae and Water ‘Clocks’	130
9.2.2	Other Timekeeping Devices	131
9.3	The First Mechanical Clocks	132
9.4	Pendulum Clocks	133
9.4.1	Galileo	133
9.4.2	Huygens	134
9.4.3	Pendulum Clock Developments	136
9.4.4	Chronometers	138
9.5	Quartz Crystal Clocks	139
9.6	Clock Performance	142
9.6.1	Quality (Q) Factor	142
9.6.2	Precision	143
9.6.3	Accuracy	143
9.6.4	Stability	144
	References	149
10	Microwave Atomic Clocks	151
10.1	Beyond Quartz-Crystal Oscillators	151
10.2	Physics of Atomic Clocks	152
10.3	General Structure of Atomic Clocks	154
10.4	Development of Atomic Clocks	157
10.4.1	Cesium	158
10.4.1.1	Calibration of the Cesium Frequency	160
10.4.1.2	Cesium Beam Tubes	162
10.4.1.3	Cesium Fountains	166
10.4.2	Hydrogen	169
10.4.2.1	Active Hydrogen Maser	170
10.4.2.2	Passive Hydrogen Masers	171
10.4.3	Rubidium	172
10.4.3.1	Rubidium Cells	172
10.4.3.2	Rubidium Fountains	173
10.4.3.3	Double-bulb Rubidium Maser	173
10.5	Stored Ion Clocks	174
10.5.1	Mercury	175
10.5.2	Other Ions	177
10.6	Characterizing Atomic Clocks	177
	References	178
11	Optical Atomic Standards	181
11.1	Optical Transition Frequencies	181
11.2	Optical Ion Clocks	183
11.3	Optical Neutral Atom Clocks	184

11.4	Quantum Logic Clock	184
11.5	Characterizing Optical Standards	186
	References	186
12	Definition and Role of a Second	189
12.1	The Historical Second	189
12.2	The Ephemeris Second	191
12.3	The SI Second	192
12.4	Adopting the SI Second	195
	References	197
13	International Atomic Time (TAI)	199
13.1	Constructing an Atomic Time Scale	199
13.2	History of TAI	201
13.3	Formation of TAI	207
13.3.1	EAL	208
13.3.2	Steering EAL with Primary Frequency Standards	215
13.4	Stability of TAI	216
13.5	Distribution of TAI	216
13.6	Relationship of TAI to Terrestrial Time	218
	References	222
14	Coordinated Universal Time	223
14.1	Universal Time Before 1972	223
14.2	Coordinated Universal Time After 1972	227
14.3	Leap Seconds	229
14.4	UT1	231
14.5	UTC Worldwide	231
14.6	Time Distribution	232
14.7	The Future of UTC – Leap Seconds or Not?	232
	References	233
15	Time in the Solar System	235
15.1	The Solar System	235
15.2	Pursuit of Uniformity	236
15.3	Pursuit of Accuracy	236
15.4	Time and Phenomena	237
15.4.1	Eclipses, Occultations, Transits	237
15.4.2	Sunrises and Sunsets	239
15.4.3	Moonrises and Moonsets	239
15.5	Time and Distance	240
15.5.1	Meter Definition	241
15.5.2	Radar Ranging	241
15.5.3	Laser Ranging	242
15.5.4	Navigation Systems	242

15.6	Space Mission Times	246
15.6.1	Doppler Effect	246
15.7	Proper Times at Planets	246
15.8	Pulsars – An Independent Source of Time?	247
	References	248
16	Time and Frequency Transfer	249
16.1	Historical Transfer Techniques	249
16.2	Time and Frequency Dissemination Modeling	250
16.2.1	Propagation Effects	250
16.2.2	Calibration	251
16.2.3	Relativistic Effects	251
16.2.3.1	Clock Transport in a Rotating Reference Frame	252
16.2.3.2	Nonrotating Local Inertial Reference Frame	252
16.2.3.3	Electromagnetic Signals Transfer in a Rotating Reference Frame	253
16.2.3.4	Electromagnetic Signals Transfer in a Nonrotating Local Inertial Reference Frame	253
16.3	Time and Frequency Dissemination Systems	254
16.3.1	Coaxial Cable	254
16.3.2	Telephone	254
16.3.3	Optical Fiber	254
16.3.4	Microwave Links	254
16.3.5	Television Broadcast	255
16.3.6	Internet	255
16.3.7	High-Frequency Radio Signals	255
16.3.8	Low-Frequency Broadcast Radio Signals	256
16.3.9	Low-Frequency Navigation Signals	257
16.3.10	Navigation Satellite Broadcast Signals	258
16.3.10.1	Global Positioning System	258
16.3.10.2	GLONASS	258
16.3.10.3	GALILEO	259
16.3.10.4	Beidou/Compass	259
16.3.11	Navigation Satellite Carrier Phase	259
16.3.12	Two-Way Satellite Time and Frequency Transfer	259
	References	260
17	Modern Earth Orientation	263
17.1	Terrestrial to Celestial Reference Systems	263
17.2	Determination of Earth Orientation Parameters	264
17.2.1	Very Long Baseline Interferometry (VLBI)	265
17.2.2	Global Positioning System (GPS)	270
17.2.3	Satellite Laser Ranging (SLR)	272
17.2.4	Doppler Orbit Determination and Radiopositioning Integrated on Satellite (DORIS)	276

17.2.5	Geophysical Modeling	276
17.3	Earth Orientation Data	279
	References	280
18	International Activities	281
18.1	Time and International Activities	281
18.2	Treaty of the Meter	281
18.2.1	General Conference on Weights and Measures (CGPM)	282
18.2.2	International Committee on Weights and Measures (CIPM)	282
18.2.3	BIPM	283
18.3	Scientific Unions	283
18.3.1	International Astronomical Union (IAU)	284
18.3.2	International Union of Geodesy and Geophysics (IUGG)	285
18.3.3	International Telecommunications Union (ITU)	286
18.4	Service Organizations	288
18.4.1	International Earth Rotation and Reference Systems Service (IERS)	290
18.4.2	International VLBI Service for Geodesy and Astrometry (IVS)	293
18.4.3	International Laser Ranging Service (ILRS)	293
18.4.4	International GNSS (Global Navigational Satellite Service) Service (IGS)	294
18.4.5	International DORIS Service (IDS)	295
	References	296
19	Time Applications	297
19.1	Time Enables the Infrastructure	297
19.2	Positioning and Navigation Services	297
19.3	Communications	298
19.4	Power Grid	300
19.5	Banking and Finance	301
19.6	Emergency Services	301
19.7	Water Flow	301
19.8	Summary	301
	Reference	304
20	Future of Timekeeping	305
20.1	Future Needs for Time	305
20.2	Modeling the Earth's Rotation	306
20.3	Clocks of the Future	307
20.4	Future Time Scales	307
20.5	Future Time Distribution	308

Glossary	311
-----------------	------------

Index	337
--------------	------------

Preface

Everyday use of time in one form or another is a common experience for everyone throughout their lives. The availability of a means to measure the passage of time with the required accuracy is taken for granted. However, the concepts on which those time scales are based and the requirements for accuracy in many applications can be both sophisticated and complex. Time is thus not a simple subject.

During the twentieth century the variability of the Earth's rotational speed was established. The basis for time that had served for so many centuries was no longer adequate to meet the more demanding needs. A search for the definition and introduction of a uniform second and a uniform time scale followed, leading to the introduction of Ephemeris Time based on the orbital motions of solar system bodies. At the same time, atomic clocks were being developed, which offered a more convenient and accurate basis for time. Time measurement progressed from scales based on the rotation of the Earth to those based on atomic physics. In addition, improvements in the accuracy of predicting planetary positions required the introduction of dynamical time scales that recognized the role of general relativity in time keeping. Over the same period, the accuracy of time keeping and time transfer improved significantly, and requirements for time have become even more demanding. The atomic second quickly achieved the status of being the most accurate and fundamental unit of measurement.

Although the Earth was no longer the basis for the most precise time keeping, the demands of new technologies made it even more critical to observe, analyze, and predict the actual variations of its rotation. The motions of its rotational axis, both in space and in the Earth itself, also required a parallel effort of observations, analysis and prediction. These activities pushed the improvement of celestial and terrestrial reference frames by orders of magnitude and encouraged new developments in the study of the dynamics of the Earth including the core, mantle, atmosphere, oceans, etc. and the forces acting on it due to the Sun, Moon, and planets. These studies have then gone on to spur the further development of even more accurate methods of observation.

This book is intended to tell the story of the progress in time keeping over the past century. It begins with time solely based on the rotation of the Earth, and proceeds through the discovery of the variations in Earth rotation and motions of the Earth's pole. During that time, clocks progressed through the improvements

in mechanical clocks to the development and improvements of atomic clocks. The availability of atomic time, the routine observations of the variable Earth rotation, and the development of the theory of relativity led to the introduction of Universal Time, International Atomic Time, Coordinated Universal Time, and a family of dynamical time scales. In the process there have been a number of scientific discoveries, significant improvements in accuracy, the development of new applications of accurate time, and the growth of the scientific field of Earth Dynamics.

A list of acronyms and a glossary are included to facilitate the use of a number of specialized terms that have appeared over the years in this field.

It is our pleasure to acknowledge and thank Ray Duncombe, Owen Gingerich, and Sergei Klioner, who read parts of the book and offered constructive suggestions for improvements. We thank Gregory Sheldon of the U. S. Naval Observatory Library for his assistance with reference details.

Dennis D. McCarthy
P. Kenneth Seidelmann

Acronyms

2MASS	2 Micron Astronomical Sky Survey
A1	USNO atomic time scale
A3	BIH atomic time scale
AAM	Atmospheric Angular Momentum
AD	Anno Domini (Common Era)
AGU	American Geophysical Union
AM	former BIH time scale
ANSI	American National Standards Institute
BC	Before Christ (Before Common Era)
BCE	Before Common Era
BCRS	Barycentric Celestial Reference System
BGI	Bureau Gravimétrique International
BIH	Bureau Internationale de l'Heure
BIPM	Bureau International des Poids et Mesures
c	speed of light
CCAUV	Consultative Committee for Acoustics, Ultrasound and Vibration
CCD	Charge Coupled Device
CCDS	Comité Consultatif pour la Définition de la Seconde
CCEM	Consultative Committee for Electricity and Magnetism
CCIF	International Telephone Consultative Committee
CCIR	International Radio Consultative Committee
CCIT	International Telegraph Consultative Committee
CCITT	International Telephone and Telegraph Consultative Committee
CCL	Consultative Committee for Length
CCM	Consultative Committee for Mass and Related Quantities
CCPR	Consultative Committee for Photometry and Radiometry
CCQM	Consultative Committee for Amount of Substance – Metrology in Chemistry
CCRI	Consultative Committee for Ionizing Radiation
CCT	Consultative Committee for Thermometry
CCTF	Consultative Committee for Time and Frequency
CCU	Consultative Committee for Units
CDMA	Code Division Multiple Access

CDS	Centre des Données Astronomiques de Strasbourg
CE	Common Era (Anno Domini)
CEO	Celestial Ephemeris Origin
CEP	Celestial Ephemeris Pole
CGPM	Conférence Générale des Poids et Mesures
CIF	Celestial Intermediate Frame
CIO	Celestial Intermediate Origin
CIP	Celestial Intermediate Pole
CIPM	Comité Internationale des Poids et Mesures
CIRF	Celestial Intermediate (or True) Reference Frame
CIRS	Celestial Intermediate Reference System
CNES	Centre National d'Etudes Spatiales
cpsd	cycles per sidereal day
CRS	Celestial Reference System
Cs	Cesium
CSNPC	China Satellite Navigation Project Center
CTRF	Conventional Terrestrial Reference Frame
CTRS	Conventional Terrestrial Reference System
DE	Development Ephemeris
DMA	Defense Mapping Agency (then NIMA, now NGA)
DORIS	Doppler Orbit Determination and Radiopositioning Integrated on Satellite
DST	Daylight Saving Time
ΔT	TT – UT1
DUT1	UT1 – UTC predicted
$\Delta UT1$	UT1 – UTC
EAL	Echelle Atomique Libre (Free Atomic Time Scale)
EDT	Eastern Daylight Time
eLORAN	enhanced LORAN
EOP	Earth Orientation Parameters
ERA	Earth Rotation Angle
ESA	European Space Agency
EST	Eastern Standard Time
ET	Ephemeris Time
ET _R	Ephemeris Time Revised
FAGS	Federation of Astronomical and Geophysical Data Analysis Services
FCN	Free Core Nutation
FDMA	Frequency Division Multiple Access
FK3	Fundamental Katalog 3
FK4	Fundamental Katalog 4
FK5	Fundamental Katalog 5
GA	Greenwich Atomic
GAST	Greenwich Apparent Sidereal Angle
GCRF	Geocentric Celestial Reference Frame
GCRS	Geocentric Celestial Reference System

GCT	Greenwich Civil Time
GEM	Goddard Earth Model
GMAT	Greenwich Mean Astronomical Time
GMST	Greenwich Mean Sidereal Time
GMT	Greenwich Mean Time
GNSS	Global Navigation Satellite Service
GPS	Global Positioning System
GRGS	Groupe de Recherche de Géodésie Spatiale
GRS	Geocentric Reference System
GRS	Geodetic Reference System
GSD	Greenwich Sidereal Date
GST	Greenwich Sidereal Time
GTRF	Geocentric True Reference Frame
HCRS	Hipparcos Catalog Reference System
HMNAO	Her Majesty's Nautical Almanac Office
HST	Hubble Space Telescope
IACS	International Association of Cryospheric Sciences
IAG	International Association for Geodesy
IAGA	International Association of Geomagnetism and Aeronomy
IAHS	International Association of Hydrological Sciences
IAMAS	International Association of Meteorology and Atmospheric Sciences
IAPSO	International Association for the Physical Sciences of the Ocean
IASPEI	International Association of Seismology and Physics of the Earth's Interior
IAU	International Astronomical Union
IAVCEI	International Association of Volcanology and Chemistry of the Earth's Interior
ICET	International Center for Earth Tides
ICRF	International Celestial Reference Frame
ICRS	International Celestial Reference System
ICSU	International Council of Science (formerly the International Council of Scientific Unions)
IERS	International Earth Rotation and Reference System Service
IGN	Institut Géographique National
IGS	International GNSS Service
IHRF	International Hipparcos Reference Frame
ILE	Improved Lunar Ephemeris
ILRS	International Laser Ranging Service
ILS	International Latitude Service
IPMS	International Polar Motion Service
ISES	International Space Environment Service
ISGI	International Service of Geomagnetic Indices
ITRF	International Terrestrial Reference Frame
ITRS	International Terrestrial Reference System
ITU-D	International Telecommunications Union-Development

ITU-ICT	International Telecommunications Union-Information and Communication Technologies
ITU-R	International Telecommunications Union-Radio
ITU-T	International Telecommunications Union-Standardization
IUGG	International Union of Geodesy and Geophysics
IVS	International VLBI Service for Geodesy and Astrometry
JD	Julian Date
JDN	Julian Day Number
JGM	Joint Gravity Model
K	degrees Kelvin
LAGEOS	Laser Geodynamics Satellite
LAT	Local Apparent Solar Time
LCT	Local Civil Time
LED	Light Emitting Diode
LHA	Local Hour Angle
LITS	Linear Ion Trap Frequency Standard
LLR	Lunar Laser Ranging
LMT	Local Mean Time
LOD	Length of Day
LORAN	Long Range Aid to Navigation
LST	Local Sidereal Time
mas	milliarcsecond
μ s	microarcsecond
MCXO	Microcomputer Controlled Crystal Oscillator
MJD	Modified Julian Day
mm	millimeter
MT	Mars Time
nas	nanoarcsecond
NASA	National Aeronautical and Space Agency
NBS	National Bureau of Standards (now NIST)
NGA	National Geospatial-intelligence Agency (formerly NIMA and DMA)
NIMA	National Imaging and Mapping Agency (now NGA)
NIST	National Institute of Standards and Technology
nm	nanometer
NPL	National Physical Laboratory
NRC	National Research Laboratory of Canada
ns	nanosecond
NTP	Network Time Protocol
OAM	Ocean Angular Momentum
OCXO	Oven Controlled Crystal Oscillator
ON	Observatoire de Neuchâtel
POSS	Palomar Optical Sky Survey
ppb	parts per billion
PPN	Parameterized Post-Newtonian
PSMSL	Permanent Service for Mean Sea Level

PTB	Physikalische Technische Bundesanstalt
PZT	Photographic Zenith Tube
QBSA	Quarterly Bulletin of Solar Activity
RGO	Royal Greenwich Observatory
rms	root mean square
s	second
SAO	Smithsonian Astrophysical Observatory
SI	Système International
SIDC	Solar Influences Data Analysis Center
SLR	Satellite Laser Ranging
SOFA	Standards of Fundamental Astronomy
SONET	Synchronous Optical Networking
SSEC	Selective Sequence Electronic Calculator
ST	Sidereal Time
STScI	Space Telescope Science Institute
TA(BIH)	BIH Atomic Time Scale
TA(k)	Atomic Time Scales from Source k
TAI	International Atomic Time
TCA	Areocentric Coordinate Time
TCB	Barycentric Coordinate Time
TCG	Geocentric Coordinate Time
TCM	Mars Coordinate Time
TCXO	Temperature-Compensated Crystal Oscillator
TDB	Barycentric Dynamical Time
TDM	Time-Division-Multiplexing
TDMA	Time Division Multiple Access
TDT	Terrestrial Dynamical Time
TEC	Total Electron Content
TEO	Terrestrial Ephemeris Origin
Teph	Barycentric Ephemeris Time
THz	Terahertz
TIO	Terrestrial Intermediate Origin
TT	Terrestrial Time
TVAR	Time Variation
TWAS	Third World Academy of Sciences
TWSFT	Two-Way Satellite Time and Frequency Transfer
UCAC	USNO CCD Astrographic Catalog
UNESCO	United Nations Educational, Scientific and Cultural Organization
URSI	International Union of Radio Science
USNAO	US Nautical Almanac Office
USNO	US Naval Observatory
UT	Universal Time
UT0	Universal Time 0
UT1	Universal Time 1
UT2	Universal Time 2

UTC	Universal Time Coordinated
VLBI	Very Long Baseline Interferometer
VLF	Very Low Frequency
WGD	World Geodetic Datum
WGD2000	World Geodetic Datum 2000
WGMS	World Glacier Monitoring Service
WGS	World Geodetic System
WRC	World Radiocommunication Conference

See Table 13.1 for acronyms for institutions contributing to TAI.

1

Time: Pre-Twentieth Century

1.1

In the Beginning

The earliest people on the planet certainly recognized the cycles of the most basic astronomical motions. The Sun and Moon rose and set each day and moved through the sky with predictable patterns, and the weather followed a cycle that seemed to be related to the movement of the Sun with respect to the stars. The units of days, months, and years naturally followed. With further observations they could visualize patterns in the stars in the sky and distinguish comets and 'stars' that 'wandered' among the others. Dramatic events such as solar and lunar eclipses were observed and in some cases recorded, predicted, and their 'meanings' interpreted. The observation and measurement of these cycles were important for daily life, religious practices, and agriculture, and they became the bases for timekeeping and calendars. Apparently, some of these observations even affected the orientations of early constructions of tombs or stone circles such as Stonehenge. Naturally, different cultures developed different customs for both keeping time and developing calendars (Aveni, 2002). Differences in the accuracy of the observations and their application led to differences in the understanding of the motions and the ability to make predictions. These improvements in knowledge and accuracy still continue to drive changes in our definition and use of time in the present and into the future.

1.2

Characterizing Time

Much has been written about the topic of time, and it is necessary to acknowledge the distinctions between the concepts of time, idealized time scales, definitions of time scales, requirements for different time scales, practical realizations of time scales, and the applications of time scales. In addition, there are time units that we can measure and also nonmeasurable time units that can only be calculated.

Traditionally, we have recognized the desirability of a kind of ‘uniform time’ with basic units that always remain the same. Isaac Newton (1686) distinguished between an idealized ‘absolute time’ and the time provided by physical measurements. The use of the word ‘uniform,’ however, implies the existence of a standard of comparison to establish that it is really uniform. The failure of observations to agree using prevailing instrumentation and models based on established philosophy or theory crafted to provide that standard has led to changes in the theories as well as the means of determining time.

The search for a practically realizable uniform time drives the hunt for an ideal unit of time and the means to access it. As measurement accuracies improve it has been, and will continue to be, necessary to develop better concepts, definitions, and practical realizations of time. Today the theory of relativity imposes further considerations regarding time in the coordinate systems of the future.

1.3

Calendars

The development of calendars varied, depending largely on religion, culture, politics and economics. Religious practices and holidays along with agricultural cycles have been defined in terms of lunar sightings, solar motion, and the appearance of the stars in the sky. Hence, calendars have been based on lunar or solar motions, or a combination of the two. Unfortunately, years, months, and days are not integral multiples of each other, and this fact has led to complications in creating calendars. For example, the year as measured by the length of time for the Sun to return to the same place along its path in the sky (ecliptic) is currently equal to 365.242 1897 days of 86 400 seconds, and the length of the month measured by the Moon’s phases is 29.530 59 days. The year cannot be composed of an integral number of months or days. Historically, the counting of years in different calendars has been based on the reigns of rulers, the lives of religious leaders, and the traditional beginnings of cultures.

Today, while a number of calendars continue in use for religious or national reasons, the Gregorian calendar, initially introduced by Pope Gregory XIII in 1582 and adopted by various countries over the next 340 years, is the calendar used internationally for civil purposes (*Explanatory Supplement to The Astronomical Ephemeris and The American Ephemeris and Nautical Almanac*, 1961). It will be of satisfactory accuracy for thousands of years to come. There are a number of books and references on calendars and the computer programs for converting between them (Richards, 1998; *Explanatory Supplement to the Astronomical Almanac*, 1992). A set of chronological ‘eras’ exists based on the various calendars, and these along with current years of various calendars are tabulated each year in *The Astronomical Phenomena*.

1.4

Astronomical Observations

Astronomical observations throughout human history have led to catalogs of star positions and theories of their motions, all for the purpose of being able to predict future phenomena. In order to catalog the positions of stars, planets, the Moon, and the Sun, it was necessary to develop a reference frame. A natural choice devised in antiquity was a set of measures based on the equator, defined by the apparent diurnal motion of the stars, and the ecliptic, which is the apparent path of the Sun in the sky. The intersection of these two circles provides a natural origin useful for making angular measurements. That origin was called the equinox, since days and nights are approximately equal when the Sun is located in that direction. Declination, measured north (positive) or south (negative) of the equator in degrees, minutes and seconds of arc up to 90 degrees, provides one angular coordinate of the direction to a celestial object. The other is the measure along the equator from the equinox, designated as right ascension and measured in hours, minutes and seconds of time up to 24 hours. The sexagesimal system of measure currently in use probably is of Babylonian origin. It was widely used for scientific purposes in antiquity and is retained today in angular measurements. Hipparchus discovered the motion of the equinox in about 129 BC from comparisons of his star positions with positions determined about 150 years earlier. This motion is called precession and is about 50" per year. Further contributions to our fundamental astronomical knowledge were made by Ptolemy and various Chinese, Mayan, Middle Eastern, and Indian astronomers. An inconvenience of this system is that the circles are in motion in space.

In the seventeenth century the need for navigation and timekeeping led to the establishment of national observatories in Paris (1667), Greenwich (1675), Berlin (1701), and St. Petersburg (1725). As these sought improvements in knowledge and accuracy from observations, star catalogs were prepared, and discoveries were made. Edmond Halley showed in 1718 that the positions of the bright stars, Aldebaran, Sirius, and Arcturus, had changed by minutes of arc from their positions in antiquity. This angular motion perpendicular to the line of sight was called proper motion. In 1728 James Bradley discovered that the Earth's orbital motion caused a change of the direction of an object by approximately 30" due to the finite speed of light, and this was called stellar aberration. Bradley detected the periodic motion of the Earth's celestial pole with respect to the stars in 1748. This motion, called nutation, can reach an amplitude of about 18". In 1781 William Herschel, in the course of his systematic observations of the sky, discovered the planetary nature of Uranus from its motion. He also discovered the solar motion toward the constellation Hercules from an analysis of proper motions. Friedrich Wilhelm Bessel, Wilhelm Struve, and Thomas Henderson independently detected in 1838–49 that star positions shifted as the observer moved in the Earth's orbit. This confirmation of the Copernican theory and measurement of stellar distances,

called parallaxes, required observations with a precision of a few tenths of an arc-second (Kovalevsky and Seidelmann, 2004).

The inclusion of these effects, along with refraction, has contributed to the improvements in accuracies over the years. Improved instrumentation, including photography, has also led to accuracy improvements in observations. Hence, there was a succession of improved star catalogs from Washington and Germany, culminating with the FK5 catalog in the 1980s. These catalogs of nearby bright stars were the basis for the reference systems defined in terms of the equinox and the equator.

Newton's law of universal gravitation demonstrated that Kepler's three laws of planetary motion were the consequences of a gravitational central force. From that development, general theories of the motions of the Sun, Moon, and planets were produced by a number of scientists in different countries. Improved accuracies of these theories, along with comparisons with steadily improving observations, led to many mathematical developments, improvements in the knowledge of astronomical quantities, and the discovery of Neptune in 1849. However, there was a lack of international agreement on astronomical quantities or ephemerides. At the end of the nineteenth century agreement was reached, with international acceptance of Newcomb's constants and solar system ephemerides, based on general theories by Newcomb and Hill, which were introduced in 1900.

The lunar theories presented a more challenging situation. Since the motion of the Moon is so rapid and the perturbations so large, the development of a theory for the motion of the Moon has presented continuing problems. Many scientists have worked on the theory, including Lalande, Peirce, Hanson, Delaunay, Hill, Newcomb, Brown, Brouwer, and Eckert. A number of different methods were employed in attempting to develop lunar theories that would successfully represent observations and predict the future motion of the Moon. In this process a 'great empirical term' was introduced, and tidal friction, variable rotation of the Earth, and the secular acceleration of the Moon were discovered. Still the understanding and accurate calculation of the ephemeris of the Moon remained a challenge.

1.5 Timekeeping

The Sun's position in the sky has always been an obvious means to keep track of time. The use of shadows of sticks and sundials cast by the Sun were natural means of telling the time. These tools could indicate the time of day by the direction of the shadow and the time of year by its length. The astrolabe first appeared in the third or second century BC and provided further improvement. It can range from a simple device for measuring the angular separation between two directions to a sophisticated device that is an analog computer and databank. Thus, it could measure the altitude of the Sun, Moon, planets, and stars, determine the hour of

day or night, the latitude of the observer, and solve other astronomical problems without numerical calculations (Fraser, 1982; Dohrn-van Rossum, 1992).

The use of controlled flows of water produced alternative means of measuring time in Egypt, India, China, and Babylonia before 1500 BC. In the third century BC water clocks were being used for scientific observations (Dohrn-van Rossum, 1992). In the 8th to 11th century AD water was used to drive mechanical wheel clocks in China. Sand clocks were introduced in the late fourteenth century AD (Dohrn-van Rossum, 1992). In the late thirteenth or early fourteenth century weight-driven mechanical clocks were developed, but their actual origin is uncertain. Initially they were more decorative than accurate, and did not have minute hands. In the seventeenth century Galileo recognized the value of the pendulum as a time-keeping device, but it was Huygens in 1656 who built the first pendulum clock. It had an accuracy of 10 seconds per day providing reasonably accurate, but not highly reliable timekeeping. It was not until the late eighteenth century that clocks were improved enough to provide reliable, accurate time (Jespersen and Fitz-Randolph, 1999). Later, pendulum clocks were improved by many refinements to provide accurate sources of time for national timekeeping and for astronomy.

The challenge of safe navigation and the means of determining longitude at sea provided strong motivation for the development of robust, accurate mechanical clocks. This challenge was met by Harrison with his H4 chronometer, made in 1759. In sea tests in 1762 it only lost 5 seconds in 81 days (Sobel, 1995).

In the 1800s, providing accurate time for civil purposes became an important function of local observatories. Typically they made transit observations of the Sun by day and stars at night to determine local solar time for their location. Hence, each locality or region had its own time based on the location of the Sun in the sky.

1.6 Time Epochs

Throughout history there have been different choices for beginning the day. The ancient Egyptians began the day at dawn, but the Babylonians and Jews chose sunset. Ancient Romans switched to midnight after first using sunrise to mark the day's beginning. Sunrise was the most common choice in Western Europe before the general use of clock time. Time was counted in units of 12 hours of day and 12 hours of night. The ancient origin of 12 hour days and nights is lost, but they were transmitted from ancient Greece, Egypt, and Babylonia. A probable developmental explanation is given by Neugebauer (1957).

Since the duration of daylight and darkness varies during the year, the length of the hours differed from day to night and during the year. For astronomical purposes these seasonal hours were replaced by a subdivision of the complete period of daylight and darkness into 24 equal and constant parts, known as equinoctial hours. Apparently Hipparchus was the first to adopt the equinoctial hours in place of the unequal and varying seasonal hours. The seasonal hours spread

throughout the Graeco-Roman world during the Hellenistic period, and until mechanical clocks became common during the Middle Ages they remained in use for civil purposes. The equinoctial hours were used only in astronomical and calendrical work or for other special purposes.

According to Pliny, Hipparchus reckoned the day from midnight, but Ptolemy reckoned it from mean noon at Alexandria, and in his *Manual Tables* divided the day and night into 24 equal hours, each hour subdivided into minutes and seconds. For astronomers and navigators observing the stars at night, it was convenient to avoid a change of days during the night, so they used days counted from noon to noon. This practice was established in astronomical tables and ephemerides within the Greek, Latin and Arabic civilization. For this reason the Julian day numbers, a continuous count of days from 4713 B.C., continue to start at noon each day. Until 1925 Greenwich Mean Time started at noon each day. In 1925 Greenwich Mean Civil Time was introduced as starting at midnight, and eventually Greenwich Mean Time came to be accepted as beginning at midnight. Thus, care must be taken when using observations from before 1930 as to what time system has been used and what is the beginning of the hour count.

1.7

Time Transfer

Without widespread and comparatively rapid commerce, the need for distribution of a common time reference had little importance. Portable sundials were adequate tools to obtain the local time. In the early fourteenth century, European communities began to install weight-driven mechanical clocks in association with bells, either in town halls or churches. These served as local time references and provided a means to notify people of the time for prayers, services, assemblies, and markets. Time balls on the top of buildings, which were visible by ships in ports, were used to signal the time on a daily basis for navigators. At the U.S. Naval Observatory (USNO) this service started in 1845. Well-calibrated portable mechanical clocks provided a means to distribute precise time. Local observatories made time available in various cities, based on their astronomical observations and mechanical clocks.

With the invention of the telegraph, time signals could be distributed over very long distances. In 1865 the U.S. Naval Observatory started sending signals at 7 a.m., noon, and 6 p.m. over the fire alarm system in Washington. That signal went to the State Department, which also had a Western Union telegraph signal, so by 1867 the time signal from USNO was transferred there and sent throughout the country. In 1869 the signals were going across country for the railroads, and by 1871 the U.S. Signal service was distributing the signal to weather stations. In 1886 synchronized clocks in public offices were kept on U.S. Naval Observatory time via the signals. Telegraph signals, in conjunction with astronomical determinations of local time, were also used to measure the time difference between

distant points and, thus, to determine their relative positions (Bartky, 2000; Dick, 2003).

1.8

Rotation of the Earth

Solar time determined from astronomical observations was the independent argument used to calculate ephemerides of the solar system until the mid-twentieth century. This was done with the understanding that solar time provided a uniform measure of time. Discrepancies between observed and calculated positions of solar system objects appeared and were most evident for the Moon, because its motion is most rapid and complex. Adams (1853) showed that the observed secular acceleration of the Moon's mean motion could not be due to gravitational perturbations. That the tides exert a retarding action on the rotation of the Earth, along with a variation in the orbital velocity of the Moon, according to the conservation of momentum, was shown by Ferrel (1864) and Delaunay (1865). The possibility of an irregular rotation rate of the Earth being the explanation for lunar residuals was considered by Newcomb (1878), but he could not find corroboration from planetary observations. The correlation of the irregularities of the motions of the inner planets and the Moon to prove the irregularity of the rotation of the Earth is described in detail in Chapter 4.

In 1765 Euler predicted that, if the axis of rotation was not coincident with the principal axis of inertia, the axis of rotation would have a circular motion with respect to the Earth's crust. Using historical transit circle observations, Seth Chandler (1891, 1892) detected this effect, called 'polar motion,' which can displace the direction of the rotation axis by angles of the order of $0.5''$. However, the observed period of 433 days did not agree with Euler's theoretical prediction of 305 days. Newcomb explained the difference as being due to the nonrigid Earth. In the 1890s the International Latitude Service (ILS), with several stations at 39 degrees north latitude, was established to make optical observations to measure polar motion. That service continued until more accurate methods were developed in the 1970s. Actually, in the celestial coordinate system, the axis of maximum moment of inertia moves around the axis of rotation with a complicated pattern made up largely of an annual component and a 14-month (Chandler) component. The amplitude of the quasi-circular motion varies between $0.05''$ and $0.25''$ in a six-year beat cycle of the 2 components.

1.9

Beginning the Twentieth Century

As the twentieth century began, official time in each country was based on pendulum clock time standards. There was no international exchange of time, although there were some accurate longitude measurements by trans-oceanographic tele-

graph signals. The international monitoring of polar motion had just begun. Variations in the Earth's rotation were suspected, but not proven. Mean solar time was based on Newcomb's theory of the Sun. Astronomical observations were improving based on photography, better instrumentation, recognition of the personal equation in observation timings, and the adoption of more accurate astronomical constants. The nineteenth century had experienced significant improvements in accuracy and knowledge, but the next century was to be even more impressive.

References

- Adams, J.C. (1853) *Phil. Trans. R. Soc. London*, CXLIII, 397–406.
- Astronomical Phenomena*, annual, U.S. Government Printing Office, Washington.
- Aveni, A. (2002) *Empires of Time*, University Press of Colorado.
- Bartky, I.R. (2000) *Selling the True Time*, Stanford University Press, Stanford, CA.
- Chandler, S.C. (1891) On the variation of latitude. *Astron. J.*, **11**, 83.
- Chandler, S.C. (1892) On the variation of latitude. *Astron. J.*, **12**, 17.
- Delaunay, C.E. (1865) *Comptes Rendus Acad. Sci.*, **61**, 1023–32.
- Dick, S.J. (2003) *Sky and Ocean Joined*, Cambridge University Press, Cambridge.
- Dohrn-van Rossum, G. (1992) *History of the Hour* (translated by T. Dunlap, 1996), University of Chicago Press, Chicago and London.
- Explanatory Supplement to the Astronomical Almanac* (ed. P. Kenneth Seidelmann), University Science Books, Mill Valley, CA, 1992.
- Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac*, Her Majesty's Stationery Office, London, 1961.
- Ferrel, W. (1864) *Proc. Am. Acad. Arts Sci.*, **VI**, 379–83.
- Fraser, J.T. (1982) *The Genesis and Evolution of Time*, University of Massachusetts Press, Amherst.
- Jespersen, J. and Fitz-Randolph, J. (1999) *From Sundials to Atomic Clocks*, Dover Publications, Inc., Mineola, NY.
- Kovalevsky, J. and Seidelmann, P.K. (2004) *Fundamentals of Astrometry*, Cambridge University Press, Cambridge.
- Neugebauer, O. (1957) *The Exact Science of Antiquity*, 2nd edn, Brown University Press, Providence.
- Newcomb, S. (1878) *Washington Observations for 1875*, Government Printing Office, Washington, DC.
- Newton, I. (1686) *Mathematical Principles of Natural Philosophy* (translated by F. Cajori, 1934), 4th printing 1960, University of California Press, Berkeley, CA.
- Richards, E.G. (1998) *Mapping Time, The Calendar and its History*, Oxford University Press, Oxford.
- Sobel, D. (1995) *Longitude*, Walker and Company, New York.

2

Solar Time

2.1

Apparent Solar Time

The position of the Sun in the sky has been used as a way to measure the passage of time from ancient times. Over the centuries this basic concept developed to the point that the angle measure, equivalent to the local hour angle of the Sun, became known as apparent solar time. This time depends on the longitude of the site in question. When the observer is on the Greenwich meridian we call it Greenwich apparent solar time, and for any other location it is called local apparent solar time. The length of a day, and therefore the second (determined by the fraction of $1/86\,400$ of the day), varies during the year, because the path of the apparent Sun is inclined with respect to the equator, and because the eccentricity of the Earth's orbit affects the rate at which the Sun appears to move in the sky. Apparent solar time is the time given by sundials and was the argument in the Nautical Almanacs and national ephemerides until the early 19th century. Historically, the accuracy of determining time was restricted to the accuracy of astronomical observations of the Sun.

2.2

Mean Solar Time

Ancient astronomers were quite aware of the fact that measuring time using the daily motion of the Sun did not result in a uniform time scale. The solution of the problem goes back at least to the time of Ptolemy, who was looking for a uniform time to construct tables of the motion of the Sun. Mean solar time uses the concept of a fictitious point on the equator of the celestial reference system defined so that in one year the Sun's motion along the ecliptic is equivalent to the motion of the fictitious point along the equator. The angle measure equivalent to the local hour angle of this point became known as mean solar time. The fictitious point is sometimes called the fictitious mean Sun, or just the mean Sun. As with apparent solar time, the mean solar time depends on the longitude of the observer's location. It

also depends on the mathematical description of the motion of the fictitious point. As mechanical clocks improved, apparent solar time became less useful, and mean solar time came into wider use. Mean solar time was introduced in almanacs in England in 1834 and in France in 1835. The concept of solar time, either mean or apparent for practical applications, relies on the uniformity of the Earth's rotation. The error of this assumption eventually led to the development of dynamical time scales and revisions of Universal Time.

The determination of the right ascension of the fictitious mean Sun was dependent on tables, or a general theory of the motion of the Sun that was available at that time. Tables of the difference between apparent and mean solar time, called the *equation of time*, could be prepared from the mathematical description of the fictitious point. If mean solar time was needed for clock comparisons, the equation of time could be applied to the observed apparent solar time.

Because of the difficulty of observing the direction of the Sun with the required precision, the determination of mean solar time, in practice, came to rely on observations of the stars. If an observer has a description of the direction to the fictitious point expressed in a conventional celestial reference system, it is possible to observe the stars in the related reference frame to determine the time. The point-like images of stars make the angle measurements much easier. The positions of stars in the reference frame were provided with respect to the equinox and equator of the system, as was the right ascension of the fictitious point. Consequently in 1900 mean solar time was defined as

$$\begin{aligned} &\text{Greenwich Hour Angle of mean equinox of date} - \\ &\text{Right Ascension of fictitious mean Sun} + 12\text{h} \end{aligned} \quad (2.1)$$

Prior to 1925 mean solar time was measured from noon, and the mean solar day beginning at noon was called the astronomical day. Mean solar time measured from noon on the Greenwich meridian was designated Greenwich Mean Time (GMT). The use of the astronomical day was discontinued at the end of the year 1924 and replaced by the civil day reckoned from midnight beginning with 1 January 1925. In 1925 January 1.0 was designated as the same instant as December 31.5 in the 1924 almanacs. The American Nautical Almanac introduced Greenwich Civil Time (GCT) and Local Civil Time (LCT) as the name for the time reckoned from midnight. Consequently, starting in 1925, the times measured from midnight were variously labeled Greenwich Civil Time (GCT) or Greenwich Mean Time. The International Astronomical Union, at the Leiden General Assembly in 1928, recommended that the designation GMT not be used, because of the different meanings it had prior to 1925, and that GCT and Universal Time (UT), which are not ambiguous, be used instead. The term Greenwich Mean Astronomical Time (GMAT) was recommended when using a day beginning at noon. Finally, the name Universal Time (UT) was introduced for time measured from Greenwich midnight, even for times before 1925. The term 'Greenwich Mean Time,' referring to time measured from midnight, continued to be used for navigation purposes and as civil time in the United

Kingdom. Now GMT is only correctly used as the civil time in the United Kingdom.

2.3

Equation of Time

The difference between mean and apparent solar time is called the equation of time. Originally the equation of time was applied to apparent solar time to obtain mean solar time, now the equation of time is commonly applied to mean solar time to obtain apparent solar time. Hence, care must be exercised concerning the sign of the equation of time. The equation of time reaches a maximum value of approximately 14 minutes about 6 February and a minimum value of about -16 minutes about 3 November, as shown in Figure 2.1.

The time-varying equation of time can also be plotted as a function of the solar declination. This curve is called the analemma.

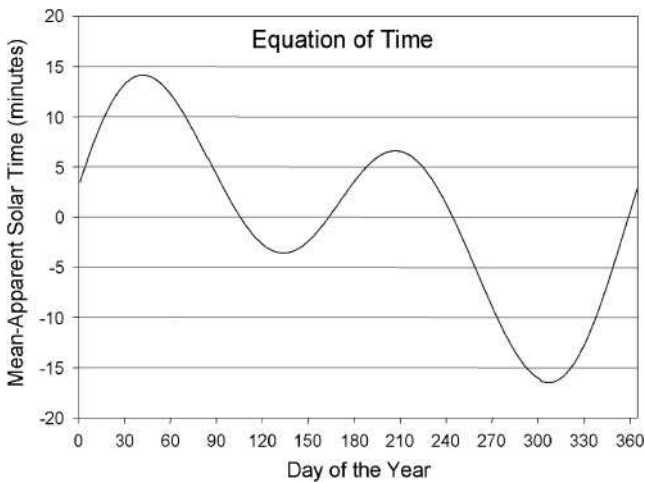


Figure 2.1 The equation of time.

2.4

Sidereal Time

The rotation of the Earth can be measured with respect to either the stars or the Sun. Because of the difficulty in making observations of the Sun, which would be required to meet the increasing demands for improved accuracy, time measured with respect to the stars became a useful measure to define mean solar time, using an adopted relationship between the two types of time measurements.

The time, equivalent to the hour angle of the equinox of the celestial frame, is called sidereal time. Since the Sun appears to move in the sky by approximately one degree each day, sidereal time is quite different from solar time. Although it is a direct and observable measure of the rotation of the Earth in the celestial system, it is not a true measure of the Earth's rotation, because the equinox is continuously moving due to precession and nutation. Apparent sidereal time is measured with respect to the true equinox, the intersection of the true equator of date and the true ecliptic of date. The true equinox of date is affected by both precession and nutation, which introduces periodic variations into apparent sidereal time. Mean sidereal time is measured with respect to the mean equinox, which is only affected by the motion due to precession. Apparent sidereal time minus mean sidereal time is the equation of the equinoxes, formerly called nutation in right ascension.

The sidereal day is the period between two consecutive transits of the equinox. Because of precession, the mean sidereal day of 24 hours of mean sidereal time is shorter than the period of rotation of the Earth by the amount of precession in right ascension in one day, which is approximately 0.0084 seconds. The sidereal day is about four minutes shorter than the solar day, because the continual motion of the Earth in its orbit during the day means that additional Earth rotation is necessary for the Sun to cross the local meridian again.

Local sidereal time depends on the longitude of the observer. Local and Greenwich sidereal times are related by

$$\text{Local sidereal time} = \text{Greenwich sidereal time} + \text{east longitude} \quad (2.2)$$

Sidereal time is usually expressed in hours, minutes, and seconds. Longitude in degrees can be converted to units of time using a conversion of one hour being equivalent to 15 degrees.

Although no longer used in practice, observations of the diurnal motions of stars do provide a measure of apparent sidereal time, since the stars' right ascensions are determined with respect to the true equinox. The equinox itself cannot be observed, and the stars are not fixed in the sky. Mean sidereal time is determined by applying the equation of equinoxes. In practice these observations require corrections for refraction, aberration, parallax, and proper motion, and allowance must also be made for the variation in the position of the meridian due to the motion of the geographic poles over the Earth's surface. Until the mid 1980s, mean solar time was derived from apparent sidereal time using a conventional expression for the right ascension of the fictitious point that defined mean solar time.

2.5

Washington Conference of 1884

The growing desire for a worldwide conventional system of longitude measure led to the International Meridian Conference held in Washington D.C., in October, 1884. The following resolutions were adopted there:

1. The Greenwich meridian would be the initial meridian for longitudes.
2. Longitudes would be measured in two directions up to 180 degrees, east longitude being positive and west longitude being negative.
3. A universal day would be adopted for all purposes for which it may be found convenient.
4. This universal day would be a mean solar day, to begin for the entire world at the moment of mean midnight of the initial meridian, coinciding with the beginning of the civil day and date of that meridian, and would be counted from zero up to twenty-four hours.

The conference expressed the hope that as soon as practicable the astronomical and nautical days would be arranged everywhere to begin at mean midnight (*Explanatory Supplement*, 1961).

The Greenwich meridian was adopted as the zero meridian, with some countries continuing to use a secondary meridian in their own country. In some cases this was because the longitude difference between Greenwich and a reference national meridian was much less accurate than local longitude differences. The astronomical and nautical day change was not made until 1925. The adoption of east positive for longitudes was not generally accepted in the west until the 1980s, when computers required the use of plus and minus numbers, rather than the letters W and E.

2.6

Newcomb's Theory of the Sun

In the 19th century there was no general agreement regarding tables of the Sun for mean solar time and for ephemerides of the Sun (see Chapter 3). However, in 1896 international agreement was reached to adopt Newcomb's *Tables of the Sun* for mean solar time (Newcomb, 1898). Thus, the basis for mean solar time became Newcomb's equation for the right ascension of the fictitious mean Sun:

$$R_{Ue} = 18\text{h } 38\text{m } 45.836\text{s} + 8640184.542\text{s}T_U + 0.0929\text{s}T_U^2, \quad (2.3)$$

where T_U is the number of Julian centuries of 36 525 days of universal time elapsed since Greenwich mean noon (12h UT) on 1900 January 0.

In his presentation of this expression Newcomb did not specify the time scale for what he called T .

Also, Newcomb's astronomical constants were adopted (Newcomb, 1895). These specified the value of precession, nutation, aberration, planetary masses, etc. that would be used until 1968, or 1984, in different cases. At the same time Newcomb and Hill's theories for the motions of the planets were adopted.

2.7

Universal Time

The term Universal Time (UT) was first recommended by the IAU in 1935 to designate mean time on the meridian of Greenwich reckoned from midnight (IAU, 1935). It now refers to a set of time scales related to the mean diurnal motion of the Sun. There has been a progression in definitions and means of determining UT, because of improvements in knowledge and accuracy. UT1 is the time scale derived from direct observations of the Earth's rotation angle in space. Astronomical observations are made to determine this angle with respect to a celestial reference system, and then it is related to mean solar time by using a mathematical expression adopted for this purpose.

Before 1984 Universal Time, was defined as 12 hours plus the Greenwich hour angle of a point on the equator whose right ascension, measured from the mean equinox of date, is given by Eq. (2.3). The measure of universal time at T_U , in hours, minutes, and seconds, was then

$$12^h + \text{Greenwich hour angle of the mean equinox of date} - R_U \quad (2.4)$$

As with sidereal time there are local mean solar times related to universal time by

$$\text{Local mean time} = \text{universal time} + \text{east longitude (in time units)} \quad (2.5)$$

Local observations of the Earth's rotation angle made by timing star transits are affected by the motion of the pole of the Earth's rotation axis over the surface. Consequently these observations provide a local measure of Universal Time known as UT0, which is a designation no longer in common use. UT1 is obtained by correcting the UT0 observations for the motion of the pole at the observing site.

UT2 corrects UT1 for the annual seasonal variation in the Earth's rotational speed by applying a conventional mathematical expression to UT1. The correction accounts for the fact that the Earth rotates more slowly in the spring and faster in the fall. UT2 is a version of universal time that was used as UT in the 1950s to 1970s, but it is rarely used in practice today. The difference between UT2 and UT1 is a conventional expression given in McCarthy, 1991.

$$UT2 = UT1 + 0.022s \sin 2\pi t_B - 0.012s \cos 2\pi t_B - 0.006s \sin 4\pi t_B + 0.007s \cos 4\pi t_B, \quad (2.6)$$

where t_B is the fraction of the Besselian year and is given by

$$t_B = \text{mod} \left[2000 + \frac{(\text{Julian Date} - 2451544.533)}{365.2422}, 1 \right]. \quad (2.7)$$

The Greenwich hour angle of the mean equinox of date is, by definition, the Greenwich mean sidereal time. At 12^h U.T. the Greenwich mean sidereal time will be R_U , which may be described as the mean sidereal time of 12^h U.T. This distinguishes it from the right ascension of the fictitious mean Sun. This led to the expression which was used from 1900 to 1984;

$$\text{GMST of } 0^{\text{h}} \text{ UT} = 6^{\text{h}} 38^{\text{m}} 45.836\text{s} + 8640184.542\text{s} T_{\text{U}} + 0.0929\text{s} T_{\text{U}}^2, \quad (2.8)$$

where T_{U} takes successive values at a uniform interval of $1/36525$ from 1900.0.

The formula was revised when improved values of astronomical constants were introduced in 1984, and the expression by Aoki *et al.* (1982) was introduced. Greenwich Mean Sidereal Time (GMST) was then related to UT1 by:

$$\begin{aligned} \text{GMST1 of } 0^{\text{h}} \text{ UT1} = & 24110.54841\text{s} + 8640184.812866\text{s} T_{\text{U}} \\ & + 0.093104\text{s} T_{\text{U}}^2 - 6.2 \times 10^{-6}\text{s} T_{\text{U}}^3, \end{aligned} \quad (2.9)$$

where

$$\begin{aligned} T_{\text{U}} = & d_{\text{U}}/36525, \quad d_{\text{U}} \text{ being the number of days of Universal Time since JD} \\ & 2451545.0 \text{ UT1 (2000 January 1, 12h UT1), taking on the values of } \pm 0.5, \\ & \pm 1.5, \pm 2.5, \pm 3.5 \dots \end{aligned}$$

The equation is consistent with the position and motion of the equinox specified by the IAU 1976 System of Astronomical Constants, the 1980 IAU Theory of Nutation, and the positions and proper motions of stars in the FK5 catalog. This equation, which is dependent on the value of the constant of precession and on variations introduced through the time scale, was considered to be the definition of UT1 until 2003.

In 2000 the XXIVth General Assembly of the International Astronomical Union (IAU) recommended the use of the ‘nonrotating origin’ (Guinot, 1979) both in the Geocentric Celestial Reference System (GCRS) and the International Terrestrial Reference System (ITRS). These origins are now called the Celestial Intermediate Origin (CIO) and the Terrestrial Intermediate Origin (TIO), respectively. The Earth Rotation Angle (ERA) is then the angle measured along the equator of the Celestial Intermediate Pole (CIP) between the CIO and the TIO. The IAU also recommended that UT1 be linearly proportional to the ERA. As a result, the Earth Rotation Angle, θ , is linked to UT1 by the conventional relationship:

$$\theta(t_u) = 2\pi \left(\begin{aligned} & \text{UT1 Julian Day Number elapsed since 2451545.0} \\ & + 0.7790572732640 + 0.00273781191135448t_u \end{aligned} \right), \quad (2.10)$$

where t_u represents the fractional part of the Julian Date.

Equation 2.10 is strictly based on the Earth’s rotation, since the ICRS is a fixed reference system (see Chapters 5 and 7). The new equation is a linear relationship between the Earth Rotation Angle (θ) and UT1, which continues the phase and rate of UT1 (Capitaine, Guinot and McCarthy, 2000). While Universal Time is no longer 12^{h} + the Greenwich hour angle of the fictitious mean Sun, it is sufficiently close compared to the difference between the mean and true Sun to continue to refer to it as mean solar time, as in the past. However, Universal Time is not precisely mean solar time.

Now, UT1 is observed using Very Long Baseline Interferometry (VLBI) measurements of selected radio point sources, mostly quasars, and interpolated by the

tracking of GPS satellites. Strictly speaking, because of the motion of satellite orbital nodes in space, VLBI provides the only rigorous determination of UT1.

Although Universal time (UT) is commonly used to mean UT1, it is sometimes used to refer to Coordinated Universal Time (see Section 2.8), and care should be exercised to make sure of the specific usage.

2.8

Coordinated Universal Time (UTC)

In the mid 1940s electronic clocks began to be used to broadcast time signals. These devices kept time with a uniform rate and were adjusted as needed to keep pace with time determined astronomically. Most time signals were adjusted in rate and offsets to match the UT2 that was determined from star transits. Each country, or organization, broadcasted their own time scale.

On 1 January 1960 the United Kingdom and the United States began coordinating adjustments made to their time scales. The resulting time scale began to be called ‘Coordinated Universal Time.’ Timing laboratories from other countries also began to participate over time, and in 1961 the Bureau International de l’Heure at Paris Observatory began to coordinate the process internationally. In 1965 the IAU officially approved the name ‘Coordinated Universal Time’ with the abbreviation UTC.

The current system of Coordinated Universal Time can be traced back to a meeting of the International Union of Radio Science (URSI) in 1966 when participants noted the need for a uniform atomic frequency. At the 1967 meeting of URSI, participants agreed that all adjustments to atomic time should be eliminated, and that UT2 information could be distributed in tables or in radio transmissions. In May, 1968, the idea of the current practice of introducing leap second adjustments in the UTC time scale was introduced independently by Louis Essen and Gernot Winkler at a meeting of a commission organized by the International Committee for Weights and Measures (CIPM) to discuss the issue. In that same year, International Radio Consultative Committee (CCIR), Study Group 7, meeting in Boulder, Colorado, discussed possible changes in the definition of UTC. They formed an ‘Interim Working Party’ to provide proposals for a possible new definition of UTC. The options considered were (i) steps in UTC of 0.1 or 0.2 seconds to keep UTC close to UT2, (ii) replacing UTC with a time scale with no adjustments, and (iii) one-second adjustments.

In 1970 the CCIR approved proposals at its XIIth Plenary Assembly in New Delhi that provide the current definition of UTC. It specified that (i) radio carrier frequencies and time intervals should correspond to the atomic second based on the cesium atom; (ii) step adjustments should be exactly one second to maintain approximate agreement with UT; and (iii) standard time signals should contain information on the difference between UTC and UT. The new system began on 1 January 1972. The difference between UT1 and UTC is provided

routinely by the International Earth Rotation and Reference Systems Service (IERS) based on astronomical observations of quasars using Very Long Baseline Interferometry.

2.9

Greenwich Mean Time (GMT)

The use of the term ‘Greenwich Mean Time’ or its abbreviation ‘GMT’ remains a source of confusion today. As discussed in Section 2.2, the mean solar time on the Greenwich meridian reckoned from noon was designated originally as Greenwich Mean Time (GMT), and the mean solar day beginning at noon, twelve hours after midnight at the beginning of the same civil day, was known as the astronomical day. From 1780 to 1833, GMT in the British Nautical Almanac was based on apparent solar time. Starting in 1834 mean solar time was the basis of the tabulations in the Nautical Almanac. The International Meridian Conference in Washington in 1884 gave special significance to GMT, as the meridian of Greenwich was designated as the zero meridian and the origin of time zones around the world.

In 1928 the IAU recommended that GMT not be used and that GCT and UT be used instead (IAU, 1928). The name Greenwich Mean Astronomical Time (GMAT) was recommended when reckoning time from noon for any dates after 1925. This confusion was modified when, in 1935, in Paris the IAU recommended that the use of GCT be discontinued, and that Universal Time (Temps Universel, in French, or Weltzeit in German) be adopted for international use. A change could not be made in the British publications, and none was made in the American Nautical Almanac. However, beginning with the 1939 volume of the American Ephemeris, the double designation ‘Universal Time or Greenwich Civil Time’ was introduced. It was still called Greenwich Mean Time in the navigation publications in English-speaking countries. Greenwich Mean Astronomical Time (GMAT) was introduced in some cases for time reckoned from noon. The astronomical almanacs ceased to use GMT after 1960; however, the navigational almanacs continued to use it as UT1. Since 1976 the IAU urged that the term GMT should be replaced by UT0, UT1, UT2, or UTC, as appropriate.

GMT was always compulsory in official British publications. The International Radio Consultative Committee (CCIR) referred to GMT for radio broadcasts of time signals, and through the 1950s GMT was used by radio time signals as being equivalent to UT2.

GMT is still used as the official time scale of the United Kingdom and in some communications systems as UTC. During the summer, GMT can be used in the United Kingdom as the name for daylight saving time, so it is UTC plus one hour. So, since 1925 the use of GMT has had different meanings and its use can be very confusing. Thus, it is recommended that for any precise timekeeping, the name GMT should not be used.

2.10

Tropical Year

Observations of the Sun are not only used to provide solar time of day; they also provide the length of the year. The tropical year is conceptually the time between successive passages of the Sun through the same point on the ecliptic. The currently accepted definition is the time for the Sun's mean longitude to increase by 360 degrees (Danjon, 1959; Meeus and Savoie, 1992). The name comes from the return of the Sun to the same tropics (i.e., tropic of Cancer or Tropic of Capricorn). The use of the word 'tropic' is derived from the Greek *tropikós* referring to 'turn'. This definition of the tropical year differs from a definition as the period of time between equinoxes. Actually, the times between the vernal equinoxes, the autumnal equinoxes, and the two solstices are different. The values for these periods for years 0 and 2000 are (Meeus and Savoie, 1992) shown in Table 2.1. It is the average of these period values that gives the value for the mean tropical year.

Table 2.1 Length of the year determined by length of time between successive passages of the Sun through the equinoxes and solstices.

	Year 0	Year 2000
March equinoxes	365.242 137 days	365.242 374 days
June solstices	365.241 726 days	365.241 626 days
September equinoxes	365.242 496 days	365.242 202 days
December solstices	365.242 883 days	365.242 740 days

The tropical year can vary by several minutes from year to year because of the motion of the Earth's perihelion, the secular increase in the rate of precession, and the periodic actions of the Moon and planets on the Earth's orbit. Averaging over time gives us a specified value for a mean tropical year. The precession rate is increasing, so the length of the tropical year is decreasing by 0.53 second per century.

The conventionally accepted value of the tropical year is 365.242 189 7 days, or 365 days 5 hours, 48 minutes and 45.19 seconds. An accurate expression for calculating its length in days in the distant past (Laskar, 1986) is

$$365.2421896698 - 0.00000615359T - 7.29 \times 10^{-10}T^2 + 2.64 \times 10^{-10}T^3, \quad (2.11)$$

where T is in Julian centuries of 36 525 days measured from 2000 January 1 Terrestrial Time (TT).

Different ephemerides will give different values for the expression for the length of the tropical year.

2.11

Besselian Year

The Besselian Solar Year is measured from the instant when the right ascension of the fictitious mean Sun, affected by aberration and measured from the mean equinox, is 18h 40m. This instant always occurs near the beginning of the calendar year. The Besselian year is shorter than the tropical year by $0.148 T$ seconds, where T is measured in centuries after 1900. This difference is due to the excess of the acceleration of the right ascension of the fictitious mean Sun over the mean longitude of the Sun (*Explanatory Supplement*, 1961). The Besselian year was used as the time measure for standard star catalogs prior to 1984. The epochs 1900.0 and 1950.0, for example, were on the Besselian year and now are designated B1900.0 and B1950.0.

2.12

Reference System

The observation of time astronomically requires definitions of terrestrial and celestial reference systems. Newcomb's Tables of the Sun, astronomical constants, the definitions of Mean Solar Time, the Tropical year, the Besselian year and the star catalog equinox provided the reference system from 1900 to 1984. When new constants and ephemerides were adopted in 1976, and introduced in 1984, Newcomb's values were replaced and the values of the tropical and Besselian year were improved with new ephemerides. Hence, the Julian year of 365.25 days was adopted for determining standard epochs. The notations of B for Besselian and J for Julian were introduced to differentiate between the two types of epochs. The standard epoch 2000.0 was introduced as a Julian epoch and designated as J2000.0. The different epochs are specified in the *Explanatory Supplements* (1961, 1992).

The reference frames were based on fundamental star catalogs of nearby bright stars. These catalogs were formed from observations, primarily made with transit circles, and included an American Series of Fundamental Catalogs starting with Newcomb and continuing to the N30 catalog as well as a series of German fundamental catalogs named FK, NFK, FK3, FK4, and FK5 (Eichhorn, 1974). The IAU recommended the use of the FK3 in 1938. It was replaced by the FK4 when it became available in 1963 and the FK5 in 1985.

The International Celestial Reference System (ICRS) was introduced in 1992 and is now maintained by the IERS and realized in practice by the directions to a set of quasars in the radio wavelengths and the directions to stars in the Hipparcos Catalog in the visual wavelengths. Similarly the IERS also maintains the International Terrestrial Reference System (ITRS) that is realized by the adopted positions of a large number of observing sites on the Earth's surface (see www.iers.org).

2.13

Time Zones

A worldwide system of standard time zones, based on increments of 15 degrees in longitude, provides the basis for local civil times that are related loosely to solar time. Time zones were first proposed in 1870 by Charles F. Dowd (Dowd, 1930) to regulate time for railroads in the United States. For political and geographical reasons the time zones are not necessarily uniform longitude strips, 15 degrees wide, running from the north to the south pole. Rather the zone boundaries are set by individual countries and usually follow country, state, or province boundaries. Some countries also introduce time zones using fractions of an hour. The zones are designated by letters, with Z specifying the Greenwich-centered zone, the zone to the east being A, progressing to the east to zone M, which specifies the 180 degree east zone. The zone for 15 west is designated N, and then the designations continue alphabetically until 180 west is designated Y. A map of the time zones is given in Figure 2.2.

2.14

Daylight Saving Time

Benjamin Franklin first suggested the idea of advancing the standard time by one hour to increase the number of daylight hours in the evening as a means of reducing energy consumption, but it took until the First World War for the idea to be adopted. Since then there has been an inconsistent history of adoption by states and countries, of the dates for the beginning and ending times, and of the political support or opposition, by locations, both in the United States and in other countries. There is also an inconsistency as to how it is designated. In some cases the names do not change, only the hours have shifted by an hour. In other cases, the word standard is replaced by Daylight, so for example, in the United States it is called informally Eastern Daylight Time in the summer and Eastern Standard Time in the winter. Daylight Saving Time is sometimes called fast, advanced, or summer time. According to the United States code, however, both times are referred to formally as standard time.

The issues concerning Daylight Saving Time continue to be matters of energy consumption, safety, airline schedule changes, agriculture, and personal living styles. One current problem, with different countries changing times on different dates, is that international airlines have to change schedules each time there is a time change to satisfy airport curfew rules.

Starting in 2007 the United States changed the daylight saving periods, which are now from the second Sunday in March to the first Sunday in November, the changes taking place at 2:00 a.m. Most of Europe goes on Advanced Time on the last Sunday in March at 2:00 a.m. local time, ending it on the last Sunday in October.

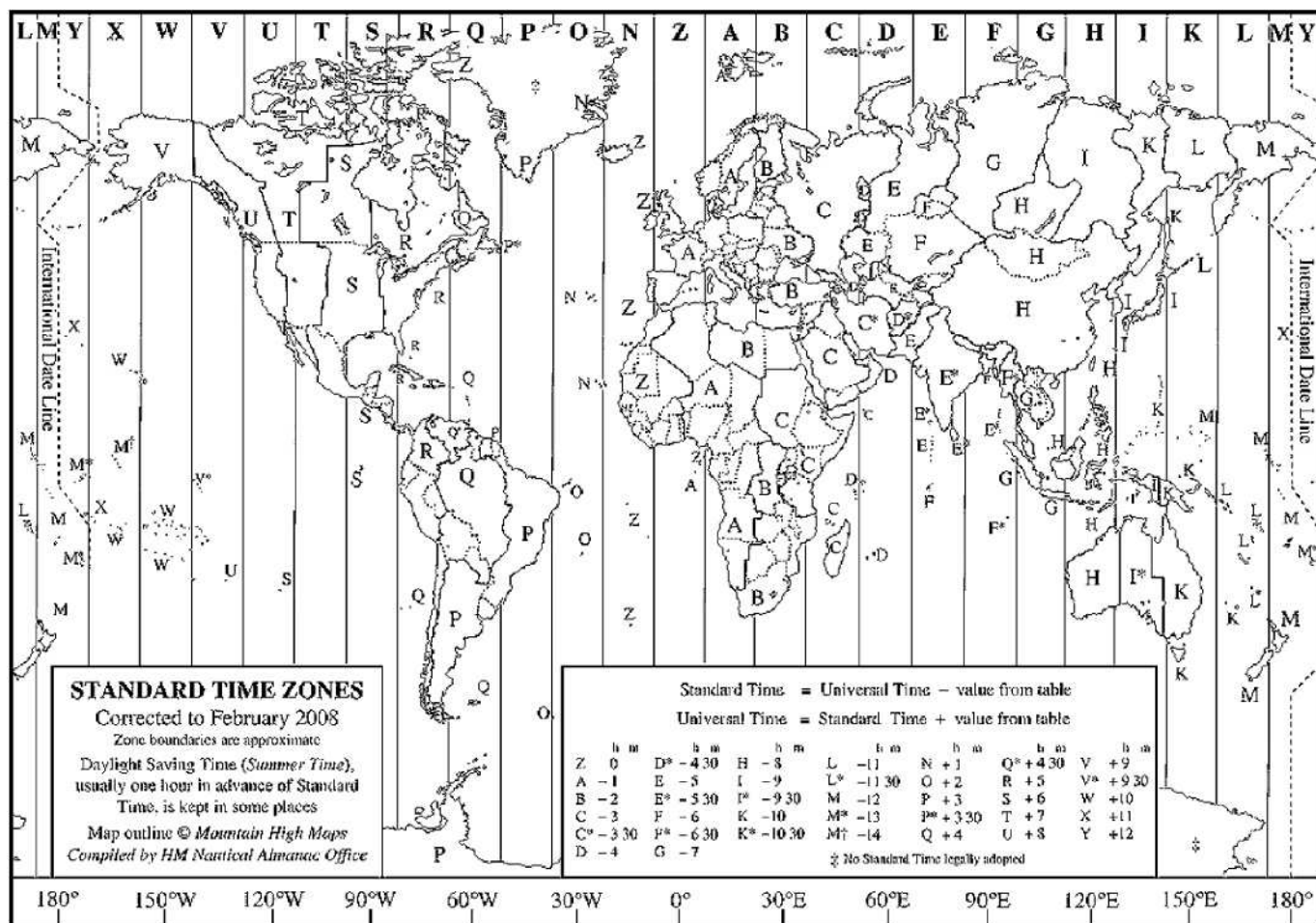


Figure 2.2 Time zone map.

References

- Aoki, S., Guinot, B., Kaplan, G.H., Kinoshita, H., McCarthy, D.D. and Seidelmann, P.K. (1982) The new definition of universal time. *Astron. Astrophys.*, **105**, 359–61.
- Capitaine, N., Guinot, B. and McCarthy, D.D. (2000) Definition of the celestial ephemeris origin and of UT1 in the international celestial reference frame. *Astron. Astrophys.*, **355**, 398–405.
- Danjon, A. (1959) *Astronomie Generale*, J. & R. Sennac, Paris.
- Dowd, C.N. (1930) Dowd, C.F. A.M., *PhD, A Narrative of His Services in Originating and Promoting the System of Standard Time*, Knickerbocker Press, New York.
- Eichhorn, H. (1974) *Astronomy of Star Positions*, Frederick Ungar Publishing Co, New York.
- Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac*, Her Majesty's Stationery Office, London, 1961.
- Explanatory Supplement to the Astronomical Almanac* (ed. P. Kenneth Seidelmann), University Science Books, Mill Valley, CA, 1992.
- Guinot, B. (1979) Basic problems in the kinematics of the rotation of the earth, in *Time and the Earth's Rotation, IAU Symp 82* (eds D.D. McCarthy and J.D.H. Pilkington), D. Reidel Publ. Co., Dordrecht, p. 7.
- IAU (1928) *Trans. IAU*, III, 224, 300.
- IAU (1935) *Trans. IAU*, V, 29–30, 286, 369.
- Laskar, J. (1986) Secular terms of classical planetary theories using the results of general relativity. *Astron. Astrophys.*, **157**, 59–70.
- McCarthy, D.D. (1991) Astronomical time. *IEEE Proc.*, **79**, 915–20.
- Meeus, J. and Savoie, D. (1992) The history of the tropical year. *J. Br. Astron. Assoc.*, **102**, 40–2.
- Newcomb, S. (1895) *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*, Government Printing Office, Washington.
- Newcomb, S. (1898) Tables of the motion of the earth on its axis around the sun. *Astronomical Papers of the AENA*, Washington.

3

Ephemerides

3.1

Ephemerides and Time

Historically, solar system positions have been made available in various formats. Some would argue that prehistoric stone circles indicate primitive knowledge of the motions of the solar system bodies. From early times tables were prepared as a way of determining the positions of the Sun, Moon, and planets. By entering these tables with a given date and performing some arithmetical operations, the position of the body could be determined for the specified epoch. The primary uses of the tables were for astrology or determining astronomical positions for specific times. Thus, it did not make sense to compute a lot of positions when only a few were needed. This lack of early ephemerides may explain the lack of improvements in astronomical positions over the centuries. It was not until the availability of the printing press that ephemerides were computed and printed (Gingerich, 2007).

The term ‘ephemerides,’ the plural of the word ‘ephemeris’ refers to tables of the positions of celestial bodies listed at regular time intervals. With improved mathematical capabilities, theories of motion were developed, so the positions of the solar system bodies could be determined from algebraic and trigonometric expressions. Later, numerical integrations were used to provide ephemerides. Currently, ephemerides are generally available in printed or electronic formats.

The time scales for the tables, theories, and ephemerides are the independent variables and it is desirable that they be uniform in rate. Through the years they have changed in order to meet the prevailing understanding of time as well as the accuracy of the observations of the solar system positions.

3.2

Before Kepler and Newton

In the earliest times observers noticed, in addition to the Sun and Moon, the bright planets, Mercury, Venus, Mars, Jupiter, and Saturn, moving among the

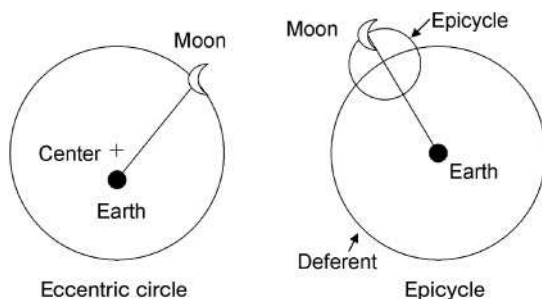


Figure 3.1 Geometric models of planetary motion.

stars. These were called wanderers, which in Greek is the derivation of the word planet. The earliest observers of the sky most likely concluded that the Earth was stationary and at the center of the universe, and this was sufficient to enable them to make predictions of eclipses and seasons. In about the 5th century BC it appears that Babylonian astronomers were developing arithmetical tables for the Sun and Moon (Neugebauer, 1969). A century later, the world system of Aristotle (384–322 BC) incorporated a series of geocentric spheres borrowed from the speculations of earlier Greek philosophers to describe the motions of celestial bodies (Dreyer, 1953). The geometric basis was a Greek innovation. Aristarchus of Samos in about 250 BC developed an earlier Pythagorean philosophy and proposed a heliocentric view of the solar system, but this was an idea without a theory, tables, or predicted positions. Although he left no astronomical writing, we know from other writers that Apollonius around 200 BC appears to have developed a geocentric model of the solar system in which the planetary motions were described using either eccentric circular motions or epicycles (Figure 3.1). In the former the planet moves uniformly along a circle, but the Earth is eccentric (i.e., not at the center), so the apparent angular speed of the Moon or planet varies. In the latter the planet moves uniformly on a circular orbit, called an epicycle, whose center moves uniformly over a circular orbit, called a deferent, around the Earth.

Hipparchus (c. 190–120 BC) from Nicaea (Turkey) and Rhodes developed accurate, geocentric models for the motion of the Sun and Moon using Babylonian and Alexandrian observations and knowledge accumulated over the centuries. He also compiled trigonometric tables, discovered precession, and compiled the first catalog of star positions. Claudius Ptolemaeus (Ptolemy) of uncertain ancestry, who lived in Alexandria, wrote the *Mathematike Syntaxis* in about 150 AD. Commonly known as the *Almagest* after the Latin form of the title of the Arab translation, it described his geocentric model and computational methods and included tables for determining solar system positions. The *Almagest* also contained a star catalog, probably an updated version of Hipparchus' catalog, and a list of 48 constellations covering the sky visible to Ptolemy. Ptolemy later produced the so-called *Handy Tables* that provided the data to compute positions of the Sun, Moon,

and planets, the rising and setting of the stars, and to predict **eclipses** of the Sun and Moon. Later astronomical tables or '**zījes**' were modeled after these tables. **Zīj** is a generic name for Arabic astronomical books that include tabular parameters for calculating positions of astronomical bodies. Ptolemy's solar tables were based on an Earth-centered solar system using circular motions and epicycles to represent the motions. The *Almagest*, which followed the geocentric cosmology of Aristotle, was generally accepted in western culture until the sixteenth century.

Aryabhata (476–550 AD), an Indian astronomer-mathematician explicitly proposed a rotating spherical Earth to explain the diurnal motion of the stars. His work *Āryabhaṭīya* written in 499 explains a geocentric system with planetary motions, being described by epicyclic models (Sharma, 1977). In the Islamic period the *Handy Tables* became the basis of the many zījes. Then followed the *Toledan Tables* and the relatively common *Alfonsine Tables*. From the twelfth century the Arabic computational 'Tables of Toledo' met the accuracy needs of the astronomers and navigators until the time of the great explorations. In the Islamic world Al-Būnā discussed the Indian heliocentric theories of Āryabhata and others in his *Chronicles of India* in 1030 AD. Arab scholars were also familiar with ancient Greek sources (Khan, 1977).

In the thirteenth century, the *Alfonsine Tables* became a widely accepted means to obtain planetary positions. These tables are named in honor of Alfonso X of Leon and Castille who encouraged the translation of a number of older Arabic publications. The *Alfonsine Tables* also contained a set of instructions written by John of Saxony that explained their use. Nevertheless, a set of 'Resolved Tables' followed that were easier to use and largely for astrological purposes. They were a different means of getting mean motions, but the equations for eccentricity and epicycle were the same as the Alfonsine's. The fifteenth century astronomer Johannes Müller von Königsberg (1436–1476), who often used the name Regiomontanus, provided, for the first time, daily positions of planets in his *Ephemerides* published for 1474–1506. He had expressed some concerns about previous tables, which he sought to improve by constructing his own, but his ephemerides closely matched the positions from the *Alfonsine Tables*.

The developments from the fifteenth century onward can, in most cases, be attributed to specific scientists, who made major contributions to our knowledge of solar system motions. In the fifteenth century, Nicolas Copernicus (1473–1543), a mathematician, astronomer, physician, classical scholar, Catholic cleric, administrator, diplomat, and economist, proposed in *De revolutionibus orbium coelestium* a heliocentric system, where he may have drawn on the Greek and Islamic tradition of mathematics and astronomy, including the works of Nasir al-Din Tusi, Mu'ayyad al-Din al-'Urdi, and ibn al-Shatir (Saliba, 2002). The detailed models generally agree because of the geometry, which had to do with the aesthetics of uniform circular motion, but did not result in improved accuracy of positions. Copernicus did provide planetary tables in his publication, but they were limited. Erasmus Reinhold (1511–1553) sought to correct that situation by publishing the *Prutenic Tables* in 1551.

3.3

Kepler and Newton

Modern work on ephemerides begins in the sixteenth century with Tycho Brahe (1546–1601) making observations of solar system bodies with improved accuracy. Johannes Kepler (1571–1630) used these observations, primarily observations of Mars, which were the only ones accurate enough to develop his three laws of planetary motion. The first two were published in *Astronomia Nova* in 1609 and the third appeared in 1619 in his work *Harmonices Mundi*. Specifically, Kepler's laws are:

1. Planets move around the Sun in elliptical orbits with the Sun located at one focus of the ellipse.
2. As the planet moves in its orbit around the Sun, equal areas, as measured from the focus, are swept out in equal times.
3. The square of the period of the orbit is proportional to the cube of the semi-major axis of the elliptic orbit.

Kepler's *Epitome Astronomiae Copernicanae* (1618–1621 in three parts) was a widely read work on theoretical astronomy. Kepler went on to publish a star catalog and tables providing planetary positions in 1627. The *Tabulae Rudolphinae* or *Rudolphine Tables* were based on the observations of Tycho Brahe and Kepler's first two laws, and were intended to complete his work. The *Rudolphine Tables* were difficult to use because they their basis on ellipses and their use of logarithms. The tables of Phillip Lansbergen (1632) appeared a few years later and were easier to use as they were based on circular orbits, but they were much less accurate. Kepler's laws and tables became the basis for further developments in Europe in the first half of the seventeenth century (Russell, 1964).

Isaac Newton (1642–1727) developed his laws of dynamics and gravitation based, in part, on Kepler's third law. His laws of dynamics can be given as:

1. Every body perseveres in its state of rest or uniform straight line motion, unless it is compelled by some impressed force to change that state.
2. The change of motion is proportional to the motive force impressed and takes place in the same direction as the force.
3. Action is always contrary and equal to reaction.

Newton's law of gravitation, as given in his *Philosophiae Naturalis Principia Mathematica* published in 1687, became the basis for Newtonian mechanics. It states that the gravitational force, f , acting between two bodies of mass, m and M , is proportional to the product of the masses and inversely proportional to the square of the distance, r , between them. This can be written as an equation in vector form

$$\mathbf{f} = \frac{GMm}{|r|^3} \mathbf{r}, \quad (3.1)$$

where G is the gravitational constant.

From this law of gravitation, ephemerides of the Sun, Moon and planets could be calculated using the available computational capabilities (Szebehely and Mark, 1998). Newton's *Principia* did not lead directly to any ephemerides, but it did influence Mayer's lunar tables and Halley's tables (Halley, 1693). The Copernican system was establishing itself, the telescope was bringing new knowledge of the skies, and the ellipse was replacing the epicycle and eccentric circle.

3.4

Tables, General Theories, and Ephemerides

In the days of computation by paper and pencil, tables of logarithms, and mental arithmetic, the calculation of astronomical positions and ephemerides was done from general theories and tables. The general theories used mean elements for the unperturbed, or Keplerian, motion of the planets, and trigonometric expressions for the periodic perturbations of one body by another. The tables gave evaluations of the trigonometric terms tabulated at time intervals small enough for the accuracy sought for the ephemerides, so the trigonometric functions did not have to be evaluated repeatedly. General theories and tables were prepared for the Sun, Moon, and planets using this process. The accuracy of the tables was limited by the accuracy of the orbital elements and the number of trigonometric terms included. The observations were then compared to the predicted positions of the ephemerides to determine improvements to the general theories and tables and to the values of the orbital elements. In this process the masses of the planets and values of astronomical constants were also determined.

In the seventeenth and eighteenth centuries ephemerides were computed by a small number of astronomers or astrologers. Kepler began one series of ephemerides in 1617 and continued until 1637, when Lorenz Eichstadt (1596–1660) continued the series in Gdansk. Johann Hecker computed ephemerides from 1666 through 1680. Then the astronomers at Paris Observatory initiated the *Connaissance des Temps* in 1679 to continue the tabular data annually. There was a competitive activity concerning the French publication, but it continued over the years with varying accuracies (Gingerich and Welther, 1983). Other contributions were made by Andrea Argoli (1570–1657) and Francesco Montebruni and his successors. Ephemerides were compiled by Eustachio Manfredi (1674–1739) in Bologna.

Thomas Streete (1621–1689) wrote *Astronomia Carolina: a New Theorie of Coelestial Motions* (1661) on computational astronomy, including tables of planetary positions and motions. This was the standard textbook into the eighteenth century and consulted by Newton, Flamsteed and Halley (Curry, 2004). Edmond Halley (1656–1742), who was personally responsible for publishing Newton's *Philosophiae Naturalis Principia Mathematica*, prepared tables of planetary positions, *Tabulae Astronomicae*, which were eventually published in 1749. He wrote an appendix for the 1710 edition of Thomas Streete's *Astronomia Carolina*, which was probably

more used than Halley's tables. Halley was also involved in improvements to the lunar tables, but the basis of the knowledge of the lunar orbit is uncertain. In 1767 the first Nautical Almanac was published in Great Britain, making use of Halley's work through the 1779 edition.

Early accurate tables of the Sun and the Moon, *Novae Tabulae motuum Solis et Lunae* 1753 were determined by Tobias Mayer (1723–1762) (Mayer, 1753), and were used for the British Nautical Almanac until 1804 (Forbes, 1980). Jérôme Lalande's (1732–1807) planetary tables (Lalande, 1792), which included mutual perturbations, represented the state of the art at the end of the eighteenth century. The *Nautical Almanac* editions from 1780 through 1804 used his tables. After that a number of different theories were used until the 1860s (*The Observatory*, 1898).

William Herschel announced the discovery of a planet, finally named Uranus, on March 13, 1781. This was the first planet discovered with a telescope and expanded the boundaries of the solar system. When the first minor planet was discovered on January 1, 1801, only a few observations were made before it was lost in the glare of the Sun, so there was a problem trying to locate it again. Calculating a prediction ephemeris from the few observations to recover the minor planet at later dates was a new problem that was solved in three months by Carl Friedrich Gauss (1777–1855). In the process he improved the mathematics of orbit prediction, which he published in *Theory of Celestial Movement* (Gauss, 1809). This introduced the Gaussian Gravitational Constant and the method of least squares, which is used to this day to minimize the effects of measurement errors. Gauss also investigated the Earth's magnetic field and developed a method to measure its horizontal intensity. He worked out the mathematical theory for separating the inner (core and crust) and outer (magnetosphere) sources of the Earth's magnetic field.

Urbain Jean Joseph LeVerrier (1811–1877) developed planetary theories, which were used for a long period, in the *Connaissance des Temps*. In the process he predicted the existence of a planet that was causing differences between observations of the position of Uranus and the positions predicted by ephemerides. His prediction for the position of the perturbing planet, which was eventually named Neptune, was independent of that of John Couch Adams, an English astronomer who was also investigating the problem. Le Verrier sent his prediction in 1846 to the German astronomer Johannes Gottfried Galle (1812–1910), who, along with his Danish assistant Heinrich Louis d'Arrest (1822–1875), found the planet on the first night of observing, in part because they had a new catalog of sufficiently faint stars. Although there was no ill will between LeVerrier and Adams, an international disagreement arose concerning whether LeVerrier or Adams should receive credit for the predicted position. In 1855 LeVerrier discovered the advance of the perihelion of Mercury, which exceeded that predicted by Newton's theory and which eventually became evidence for Einstein's theory of relativity. LeVerrier attributed the perihelion advance to a planet called Vulcan, or a second asteroid belt, which was closer to the Sun than Mercury.

The British and U.S. Nautical Almanac Offices used the tables of LeVerrier until 1900, when the tables of Newcomb and Hill were introduced for the

Sun and planets. A list of the Tables used in the British *Nautical Almanac* and *American Ephemeris and Nautical Almanac* prior to 1900 is given in the Explanatory Supplement (1961). An excellent description of the details and accuracies, including figures, showing the sources of the solar system ephemerides and the progression of accuracies from 1600 to 1800, is given by Gingerich and Welther (1983).

Many scientific developments during the 20th century both permitted and required improvements in the ephemerides of the solar system. The new developments included Einstein's Theory of Relativity, the variable rotation of the Earth, the discovery of Pluto, the need for a new uniform time scale, improved observational accuracies, including radar, laser ranging, and spacecraft observations, and improved astronomical constants. The lunar theory was improved (see Section 3.5), and the ephemerides of the five outer planets were improved with numerical integration (Eckert, Brouwer and Clemence, 1951), but the ephemerides of the other planets continued to rely on older theories until the 1980s. By 1980 the planetary ephemerides were in error by about 0.10s in right ascension and 0.3" in declination, except for Neptune and Pluto whose errors were significantly larger. In 1984 numerically integrated ephemerides by the Jet Propulsion Laboratory, designated DE200, were introduced with significantly improved accuracies. There have been many improvements in the planetary and lunar ephemerides over the last 30 years. Details concerning the numerical integrations, the equations of motion, and observational data are given by Standish and Williams (2009). The current ephemerides of the inner planets are accurate in this decade to 100–200 m, and the relative angles (Earth, Sun, planet) are accurate to 0.001". These ephemerides are based largely on radiometric ranging observations. For the outer planets the observational data are still primarily optical, so the accuracies are much poorer. The plane-of-the-sky directions are accurate to a few hundredths of an arcsecond at present, but they deteriorate significantly with time, particularly for the outermost planets. The primary limitation on the accuracy of current planetary ephemerides is the uncertainty of minor planet masses and their effect on the orbit of Mars.

3.5

Lunar Theories

The Moon is the most rapidly moving and closest celestial body to the Earth, so representing the motion of the Moon is the most complicated. Together with the Sun and the Earth, this is an example of the classical celestial mechanics three-body problem. In addition to the challenge of determining the periodic motions due to the perturbations by the Sun and the other planets, there are complicated interactions between the Earth and the Moon. Further, all the kinematics of the Earth's motion are involved in determining the lunar ephemeris.

Hipparchus studied the motion of the Moon and confirmed the values of the periods of its motion as determined by the Babylonian astronomers. The error in

their value of the synodic month was less than 0.2 second in the 4th century BC, and less than 0.1 second in Hipparchus's time. In this work Hipparchus used both the eccentric circle and epicycle models to determine relative proportions and actual sizes of the orbits.

With the founding of the Royal Greenwich Observatory in 1675, John Flamsteed (1646–1719) became Charles II's 'astronomical observator.' Now recognized as the first Astronomer Royal, he began his work at the observatory by demonstrating that the rotational speed of the Earth did not vary. Of course, the rotational speed does vary, but his clocks were not accurate enough to detect this fact. He then set about making systematic observations of the Moon to obtain the data necessary to implement the lunar distance method for determining longitude at sea (Kollerstrom and Yallop, 1995). In the course of this work Flamsteed compiled his tables of the Moon's position from 1689 to 1704.

Using Flamsteed's observations, Edmund Halley (1695) noted that the Moon's motion appeared to be gradually speeding up. This 'secular acceleration' could not be explained by gravitational theory. Although Newton had developed a lunar theory in 1702 (Gregory, 1715), Alexis Clairaut (1713–1765) computed the first full analytical theory of the Moon in 1750. Leonhard Euler (1707–1783) published lunar tables in 1746, and in 1753 provided his theory of lunar motion (Euler, 1753), which was used by Tobias Mayer in his tables. Mayer's work resulted in an award of £3000 from the British government to his widow for his contribution to the problem of longitude determination. Euler also received £300 for his theoretical work. Mayer did the first careful investigation of the librations of the Moon in 1750. In 1755 James Bradley (1693–1762), England's Astronomer Royal from 1742 until his death, compared Mayer's lunar tables to Greenwich observations and found them capable of providing the Moon's position with an accuracy of 5".

In the eighteenth century significant contributions to the lunar theory were also made by Jean le Rond d'Alembert (1717–1783) and Pierre-Simon, Marquis de Laplace (1749–1827), among others. Laplace (1786) addressed the issue of the observed secular acceleration in the Moon's motion and theorized that this was due to a secular change in the eccentricity of the Earth's orbit. In fact, it accounts for only a portion of the observed value, the rest being caused by the effect of the Earth's secular rotational deceleration due chiefly to tidal friction. In 1818 Laplace proposed that the Académie des Sciences in Paris award a prize to whoever succeeded in constructing lunar tables based solely on the law of universal gravity. In 1820 that prize was awarded to Francesco Carlini, Giovanni Plana, and Marie-Charles-Théodore de Damoiseau.

In 1846 Charles-Eugène Delaunay (1816–1872) published his work on the lunar theory, containing what became known as 'Delaunay's method' (Delaunay, 1846), and in 1865 he pointed out that problems with the lunar orbit might be due to the slowing of the Earth's rotation due to tidal friction (Delaunay, 1866). His work on the lunar theory is summarized in two volumes (1860, 1867). Other lunar theories were developed by Gustave de Pontecoulant (1840) and Peter Hansen (Brown,

1896, 1960; Cook, 1988). Hansen's tables (1857) were used generally in the almanacs after 1882. However, soon after that, corrections to Hansen's tables, determined by Simon Newcomb (1876), were gradually introduced in the published lunar ephemerides.

G. W. Hill (1838–1914) took a new approach to the lunar theory. Instead of polar coordinates referred to fixed axes, he used rectangular coordinates referred to moving axes. This method led to the use of a 'variational curve', which contains an important part of the solar perturbation, known as the 'variation.' Thus, Hill obtained differential equations in a simple algebraic form, suitable for solution in infinite series. Ernest Brown (1866–1938) used Hill's method to develop his lunar tables (Brown, 1896). Brown's Lunar Tables (Brown, 1897–1908) were used in the almanacs from 1923 to 1960, when the Improved Lunar Ephemeris (ILE) (1954), based on Brown's theory, was introduced. Observations showed that Brown's tables were better than those of Hansen, which had been used since 1857, but there was still a 10" fluctuation in the Moon's mean longitude. So, a 'great empirical term' of magnitude 10.71" and period 257 years was introduced to correct for the fluctuation. Simon Newcomb suggested that it was due to a gradual deceleration of the Earth's rate of rotation due to friction generated by the tides. In other words the Moon was not speeding up, but time, as measured by the Earth's increasingly longer day, appeared to be slowing down. Brown concluded that the Earth's rate of rotation was slowing, and that there were random, unpredictable fluctuations in the Earth's rotation with periods of 60 to 70 years. In practice this is one of the complications in separating the lunar secular acceleration, due to tidal interaction and secular change in the eccentricity of the Earth's orbit, from variations in the Earth's rotation (Brown, 1914).

The lunar theory presented many challenges to achieving accuracies to match those of the observations. In addition to the problems of developing a fully accurate gravitational theory including the Sun, Earth, and Moon interactions, the planetary perturbations, both direct and indirect effects, and the figure of Earth and Moon, all of which required extensive lists of periodic terms, there were the problems of the secular acceleration, the tidal interactions, the librational motions, limb variations, lunisolar precession and nutation, the location of the equinox, and the variable Earth rotation.

Beginning in 1960 the British and U.S. almanacs used the ILE calculated directly from Brown's lunar theory instead of from his tables. In order to obtain a gravitational ephemeris on the same measure of time as Newcomb's Tables of the Sun, the empirical term was removed from the orbital elements and the following correction was applied to the mean longitude

$$\Delta l = -8.72' - 26.74'T - 11.22'T^2, \quad (3.2)$$

where T is measured in Julian centuries from 1900 January 0.5.

Brown's theory represented the lunar main problem with accuracies of 0.01" in longitude and latitude and 0.0001" in parallax (Henrard, 1973). The perturbations

due to the Earth's flattening were less precise, about $0.07''$ in longitude and latitude. However, the planetary perturbations were shown to produce differences from numerical integrations in position by up to 500 m (Mulholland, 1969). Inaccuracies in fundamental constants may increase this error, particularly in secular and long-period terms. A probable secular error in longitude of $8'$ per century was quoted by Mulholland (1972) (Kovalevsky, 1977).

Observations of occultations, where the Moon is observed passing in front of a star, accurately measure the Moon's position with respect to the star's position in a star catalog reference system. These indicated lunar ephemeris errors of as much as $0.5''$, or 5 km, in the 1970s. The ILE was replaced in 1984 by a numerically integrated lunar ephemeris. However, the long-term trend of the lunar motion in the numerical integration was fit to that of the ILE.

3.6

The Advent of Computers

Following early attempts to ease computational burdens, including the adding machine of Blaise Pascal (1623–1662) in 1645 and the desk calculator 'step reckoner' of Gottfried Leibniz (1646–1716) in 1671, Charles Babbage (1791–1871) in the early 1800s conceived an 'Analytical Engine' to perform calculations (Dubbey, 1978). Because of lack of funds the machine was never completed, but Georg and Edvard Scheutz used the description to construct such a machine for display in London and Paris in 1854 and 1855. B.A. Gould purchased the machine for Dudley Observatory, where it was used to calculate tables of the Sun's longitude, the radius vectors of Venus and Earth, and the geocentric distance to Mars (Seidelmann, 1976).

With the introduction of punched card equipment in the 1920s, the computation of ephemerides from the tables by use of this equipment was introduced by L. J. Comrie in Britain for the determination of the lunar ephemeris (Comrie, 1925). In the United States, Wallace Eckert was hired as director of the Nautical Almanac Office of the U.S. Naval Observatory at the beginning of World War II. He introduced punched card equipment for computations and a punch card-operated typewriter to print camera-ready pages of the almanacs (Dick, 1999).

In 1948 the IBM Selective Sequence Electronic Calculator (SSEC) computed the ephemeris of the Moon directly from Brown's theory for 20 years to an accuracy of $0.01''$. The SSEC was used for a numerical integration of the Sun and five outer planets from 1753 to 2060 at 40 day intervals with an accuracy of 16 digits (Eckert, Brouwer and Clemence, 1951). These became the basis for the publication of the outer planet positions in the *American Ephemeris*.

With the capabilities of modern computers, more accurate ephemerides could be determined by numerical integrations than from general theories. So, when new astronomical constants, a new fundamental star catalog, and a new dynamical reference frame were introduced in 1984, ephemerides based on numerical integrations were also introduced (Seidelmann, 1976).

3.7

Numerical Integrations

Numerical integration of differential equations involves the use of initial conditions, specifically in the case of ephemerides, a position and velocity vector for each body. The standard method used is based on the formulation by P. H. Cowell (1870–1949) and uses an equation of motion that can incorporate all the appropriate forces along with terms for relativistic effects. The perturbations from all bodies can be included (*Explanatory Supplement to the Astronomical Almanac*, 1992). The derivatives and positions are determined from the differential equations of motion. Then difference tables, evaluated for successive time steps, are used to determine the successive positions and velocities of the bodies. Once an ephemeris is determined it is compared to observations, a least-squares solution is made, and a differential correction is made to determine new initial conditions, and the next numerical integration is started. By an iterative process, the numerical integration is fitted to the observations as well as possible. This process is computationally intensive, so that it is time consuming when done by hand but easily carried out on a computer.

3.8

Observational Data

Solar system observational data have increased in quantity and accuracy over the years. Until the 1960s the data were restricted to optical observations. These were primarily transit circle observations. The solar and lunar observations were particularly subject to systematic errors, since they differed significantly from stellar observations and involved observing relatively bright limbs of the bodies. In addition, solar observations were affected by instrumental heating. The inner planets were subject to phase effects based on their positions relative to the Sun. Lunar occultation observations provided a different type of observation. Systematic discussions of the different planetary observations were undertaken by a number of individuals, resulting in improved ephemerides. Examples are the investigations of Mercury by Clemence (1943), Venus by Duncombe (1958), Mars by Ross (1917) and Laubscher (1981), and Neptune by Jackson (1974).

There were significant improvements in the observational data due to improved methods, instruments, and scientific knowledge, particularly in about 1830 and again in about 1900. Hence, for those planets with shorter periods of revolution about the Sun, observations before 1900 were not used. For those with longer periods, like Uranus and Neptune, all available observations had to be used.

Since Pluto was not discovered until 1930, and pre-discovery observations only go back to 1914, the observations do not cover the entire 250-year period of Pluto. Also, since Pluto is only a 15th magnitude object, its observations had to be made photographically and with respect to faint reference stars whose positions were not well determined.

3.8.1

Radar Observations

Radar observations of planets began in 1966 with timing by atomic clocks. The observed delay in receiving the reflected signal depends not only on the distance of the target but it is also affected by relativistic effects, the solar corona, and the troposphere. The variations in planetary topography are responsible for the largest uncertainties in the radar measurements. For the inner solar system the radar observations dominate the determination of the ephemerides and the dynamical reference system (*Explanatory Supplement to the Astronomical Almanac*, 1992).

3.8.2

Lunar Laser Ranging

The Apollo 11, 14, and 15 missions landed retro-reflectors on the Moon at their landing sites starting in 1969. Also, there is a retro reflector on the Lunakhod 2 vehicle. Since the lunar laser ranging measurements are from an observatory fixed on the surface of the Earth to a retro-reflector fixed on the surface of the Moon, the observations of the length of time for the laser signal to be reflected from the target are dependent on the distance between the Earth and the Moon, the librations of the Moon, and all the kinematics of the Earth. So the measurements can be used to determine the orbit and the librations of the Moon, the rotation of the Earth, polar motion, precession, nutation, crustal motion, and relativistic effects.

The lunar laser ranging measurement accuracies have progressed from the centimeter level to the millimeter level. At this accuracy level the Weak and Strong Equivalence Principles can be tested with sensitivity approaching 10^{-14} . In addition to improving values of the post-Newtonian (PPN) parameters, effects of the next post-Newtonian orders (c^4) of light deflection, and other relativistic effects can be investigated (Turyshev *et al.*, 2004).

3.8.3

Spacecraft Observations

With spacecraft missions to planets, round-trip range and Doppler measurements became available. The spacecraft include missions making close passes by the planets, those in orbits around the planets, and landers on the surfaces of planets. These data provide distances to the planets, planet masses and gravity fields, and satellite information (*Explanatory Supplement to the Astronomical Almanac*, 1992).

3.9

Dynamical Reference Frame

Prior to the 1990s, astronomical reference frames were realized by the adopted positions and motions of bright stars. The equinox and mean solar time were

conventionally defined by Newcomb's Tables of the Sun, (Newcomb, 1898), and his astronomical constants completed the definition of the celestial reference systems. There was the continuing problem of the differences in the coordinate systems implied by the solar, lunar, and planetary ephemerides, and those of the star catalogs. In the 1950s discrepancies between the origins of the solar and lunar ephemerides were recognized and corrections were introduced. By the 1970s it was recognized that the equinox defined by the solar system ephemerides, called the dynamical equinox, differed from the origin of the fundamental star catalog right ascensions, called the catalog equinox, and that the star catalog origin of right ascensions varied by declination because of systematic errors in the star catalogs. Also, it was recognized that the presence of systematic errors in the proper motions of stars could cause an apparent rotation of the equinox of the star catalogs. Thus, an equinox motion correction was introduced between the FK4 (Fricke and Kopff, 1963) and FK5 (Fricke, Schwan and Lederle, 1988) star catalogs.

Prior to the 1960s the ephemerides were determined using optical observations of the positions of the Sun, Moon, and planets. With the availability of radar, lunar laser, and spacecraft observations, the ephemerides for the inner planets and the Moon began to make use of new types of observations, and the reference frame of the ephemerides grew to become independent of the star catalogs. Each new ephemeris defined its own version of the equinox by its origin of right ascension.

It was also recognized that two methods of determining the position of the dynamical equinox were being used. One method is based on a 'rotating' ecliptic related to the geometrical path of the Earth-Moon barycenter. This has been used historically since Newcomb. The other method is based on an ecliptic pole defined by the mean orbital angular momentum vector of the Earth-Moon barycenter in a Barycentric Celestial Reference System. This is an 'inertial' ecliptic, related to the Earth-Moon barycentric orbital angular momentum vector. The difference between these two methods of determining the equinox is about 0.1" (Hilton *et al.*, 2006).

Starting in 1992 a new reference system based on extragalactic radio sources was developed. This is called the International Celestial Reference System and is described in detail in later chapters.

3.10 Time Arguments

The independent variable in ephemerides is uniform time. From the time of the Hellenistic astronomer Ptolemy, the concept of mean solar time was used for this independent variable. Following the recognition of the variability of the Earth's rotation, a new time scale, Ephemeris Time, based on the Earth's orbital motion was introduced (See Chapter 5). Problems with real-time realization of Ephemeris Time soon became apparent, and atomic time became available as an alternative source of Ephemeris Time. With improvements in accuracies of observations and

ephemerides, it became necessary to include relativistic theories in the definitions of time scales, so a category of dynamical times was introduced in 1984. These time scales were refined in definitions and names in the 1990s, so there are dynamical time scales for ephemerides that are precisely related to an atomic time scale. Details concerning these developments are described in the following chapters.

3.11

Astronomical Constants

Each ephemeris requires a set of astronomical constants, either explicitly adopted for that purpose or developed with the ephemeris. These constants might include gravitational constants, planetary masses, the astronomical unit, the precession constant, nutation theory, aberration constant, and so forth. Newcomb became the first to undertake a completely new set of consistent ephemerides for the Sun, Moon, and planets, along with the determination of a complete set of astronomical constants. This collected work was adopted by the international community in 1900, including an expression for mean solar time based on his solar theory (Newcomb, 1895, 1898).

The improving accuracies of observations drove further improvements in the accuracies of the ephemerides and the astronomical constants. Newcomb's constants were partially updated by the IAU in 1968 (*Supplement to the A.E. 1968*, 1966). With the introduction of a new reference frame and new ephemerides based on numerical integration, new time scales and new astronomical constants in 1984, Newcomb's system was finally replaced to meet the requirements for better accuracies (*Explanatory Supplement to the Astronomical Almanac*, 1992). The introduction of an improved astronomical system of constants, star catalog, and ephemerides, along with new observational techniques, provided the means for further improvements in accuracies.

In current practice, the values of some of these constants may vary at the highest levels of precision depending on the application. The International Union of Geodesy and Geophysics (IUGG), the International Earth rotation and Reference system Services (IERS) and the IAU are sources for conventional values that are updated as required to be consistent with modern applications. Producers of modern ephemerides also generally publish the numerical values of the constants employed in the production of their ephemerides, and tables of constants can be found in the *Astronomical Almanac*.

3.12

Artificial Satellite Theories

With the launch of Sputnik on October 4, 1957, the computation of orbits for artificial satellites created the field of astrodynamics. This topic is essentially celes-

tial mechanics related to objects whose propulsion systems can be used to change their orbits. Artificial satellites could be launched for a variety of purposes, including communication, navigation, reconnaissance, and scientific observations. The use of artificial satellites for these purposes introduced new accuracy requirements for orbit computations and for timekeeping. Analysts can choose among geocentric rotating or nonrotating, or barycentric reference frames depending on their needs. The relativistic equations must be formulated according to the reference frame chosen. The use of satellites for time transfer improved the international system of timekeeping, and the use of accurate clocks on satellites introduced a new means of accurate navigation, specifically the Global Positioning System (GPS).

3.13

Theory of Relativity

While Einstein published his theory of relativity in 1905, the incorporation of the theory into ephemerides, time scales, and astrometry was delayed, because the accuracies of observations and theories did not require the inclusion of relativity, except for the advance of the perihelion of Mercury. The introduction of Ephemeris Time in 1952 did not take the theory of relativity into consideration. The ephemerides of the Sun, Moon, and planets were not changed to correct for relativity until the numerically integrated ephemerides were introduced in 1984. In this case the parameterized post-Newtonian (PPN) n-body metric equations of motion were introduced (Will, 1974). Planetary observations were not corrected for relativistic effects until after 1984. The dynamical time scales were introduced in 1984 to recognize relativistic effects. Atomic time scales and time transfer were specified in conformity with relativity starting in the 1970s. Subsequent chapters provide the details of these changes.

References

- | | |
|--|---|
| <p>Brown, E.W. (1860) <i>An Introductory Treatise on the Lunar Theory</i>, Dover Publications, New York.</p> <p>Brown, E.W. (1896) <i>An Introductory Treatise on the Lunar Theory</i>, Cambridge University Press, Cambridge.</p> <p>Brown, E.W. (1897–1908) <i>Memoirs of the Royal Astronomical Society</i> 1960, 48, Royal Astronomical Society, London.</p> <p>Brown, E.W. (1914) Cosmic physics. <i>Science</i>, Vol. XL, 389–401.</p> <p>Clemence, G.M. (1943) The motion of mercury, 1765–937, in <i>Astronomical Papers of the American Ephemeris and Nautical Almanac</i>, Vol. XI, U.S. Government Printing Office, Washington.</p> | <p>Comrie, L.J. (1925) The application of calculating machines to astronomical computing. <i>Pop. Astron.</i>, 33, 1–4.</p> <p>Cook, A. (1988) <i>The Motion of the Moon</i>, Adam Hilger, Bristol and Philadelphia.</p> <p>Curry, P. (2004) Streete, Thomas (1621–1689), in <i>Oxford Dictionary of National Biography</i>, Oxford University Press, Oxford.</p> <p>Delaunay, C.-E. (1846) <i>Mémoire sur une Méthode nouvelle pour la détermination du mouvement de la Lune</i>, Comptes Rend. Acad. Sci., 22.</p> <p>Delaunay, C.-E. (1860, 1867) <i>La Théorie du mouvement de la lune</i>, Monthly Notices of RAS, 21, 80.</p> |
|--|---|

- Delaunay, C.-E. (1866) *Ralentissement de la rotation de la terre*, G. Bailliere, Paris.
- Dick, S. (1999) 'History of the American Nautical Almanac Office', The Eckert and Clemence Years, 1940–1958, in *Proceedings, Nautical Almanac Office Sesquicentennial Symposium* (eds A.D. Fiala and S.J. Dick), U.S. Naval Observatory, Washington DC, pp. 35–46.
- Dreyer, J.L.E. (1953) *The History of Astronomy from Thales to Kepler*, Dover Publications.
- Dubbey, J.M. (1978) *The Mathematical Work of Charles Babbage*, Cambridge University Press, Cambridge.
- Duncombe, R.L. (1958) The motion of Venus, 1750–1949, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. XVI, U.S. Government Printing Office, Washington.
- Eckert, W.J., Brouwer, D. and Clemence, G.M. (1951) Coordinates of the five outer planets, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. XII, U.S. Government Printing Office, Washington.
- Euler, L. (1753) *Theoria Motus Lunae exhibens omnes eius inaequalitates*, Impensis Academiae Imperialis Scientiarum, St Petersburg.
- Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac*, Her Majesty's Stationery Office, London, 1961.
- Explanatory Supplement to the Astronomical Almanac* (ed. P. Kenneth Seidelmann), University Science Books, Mill Valley, CA, 1992.
- Forbes, G.E. (1980) *Tobias Mayer (1723–62), pioneer of enlightened science in Germany*, Vandenhoech & Ruprecht, Göttingen.
- Fricke, W. and Kopff, A. (1963) Fourth fundamental catalogue and supplement. *Veröff. Astron. Rechen-Inst., Heidelberg*, **10**, 11.
- Fricke, W., Schwan, H. and Lederle, T. (1988) Fifth Fundamental Catalogue (FK5). Part I. The basic fundamental stars. *Veröff. Astron. Rechen-Inst. Heidelberg*, **32**, 1–106.
- Gauss, C.F. (1809) *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solem Ambientium*, translation: *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections*, Dover Publications, August 20, 2004.
- Gingerich, O. (2007) Gutenberg's gift, in *Library and Information Services in Astronomy V: Common Challenges, Uncommon Solutions, ASP Conference Series*, Vol. 377 (eds S. Ricketts, C. Birdie and E. Isaksson), Astronomical Society of the Pacific, pp. 319–28.
- Gingerich, O. and Welther, B.L. (1983) *Planetary, Lunar, and Solar Positions, New and Full Moons, A.D. 1650–805*, Memoirs Series, Vol. 59S, American Philosophical Society.
- Gregory, D. (1972) *Astronomiae Physicae at Geometricae Elementa*, 1702, with 'Lunae Theoria Newtoniana' on pp. 332–6; reprinted 1726; translated as *The Elements of Physical and Geometrical Astronomy*, 1715, 2nd edn 1726, 2 vols, with TMM on pp. 563–71; facsimile reprint (Sources of Science, no.119) NY & London, Johnson Reprint.
- Halley, E. (1693) *Emendationes ac Notae in vetustas Albatennii Observationes Astronomicas cum restitutione Tabularum Lunisolarium ejusdem Authoris*. *Phil. Trans. R. Soc.*, **17**, 913.
- Halley, E. (1695) Some account of the ancient state of the city of Palmyra, with short remarks upon the inscriptions found there. *Phil. Trans. R. Soc.*, **19**, 160.
- Hansen, P. (1857) *Tables De La Lune, Construites D'après La Principe Newtonien De La Gravitation Universelle*, G. E. Eyre et G. Spottiswoode, London.
- Henrard, J. (1973) L'éphéméride analytique lunaire–ALE. *Ciel et Terre*, **89**, 1.
- Hilton, J.L., Capitaine, N., Chapront, J., Ferrandiz, J.M., Fienga, A., Fukushima, T., Getino, J., Mathews, P., Simon, J.-L., Soffel, M., Vondrak, J., Wallace, P. and Williams, J. (2006) Report of the International Astronomical Union Division I Working Group on precession and the ecliptic. *Celest. Mech. Dyn. Astron.*, **94**, 351–67.
- ILE (1954) *Improved Lunar Ephemeris 1952–1959*, U.S. Government Printing Office.
- Jackson, E.S. (1974) A discussion of the observations of Neptune 1846–970, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. XXII, U.S. Government Printing Office, Washington.

- Khan, M.S. (1977) ryabhata I and Al-B run . *Indian J. Hist. Sci*, **12**, 237.
- Kollerstrom, N. and Yallop, B. (1995) Flamsteed's Lunar Data 1692–5, sent to Newton. *J. Hist. Astron.*, XXVI, 237–46.
- Kovalevsky, J. (1977) Lunar orbital theory. *Phil. Trans. R. Soc. London A*, **284**, 565–71.
- Lalande, J. (1792) *Traité d'astronomie*, 2 Vols, 1764 enlarged. edn, 4 Vols, 1771–1781; 3rd edn, 3 Vols, P. Didot.
- Lansbergen, P. (1632) *Tabulae motuum coelestium perpetuae*, apud Zachariam Romanum, Middelburgi Zelandiae.
- Laplace, P.-S. (1786) Sur l'équation séculaire de la Lune. *Mém. Acad. Roy. Sci.*, 235.
- Laubscher, R.E. (1981) The motion of mars 1751–1969, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. XXII, U.S. Government Printing Office, Washington.
- Mayer, T. (1753) *Novae Tabulae motuum Solis et Lunae*, in *Commentarii Societatis Regiae Scientiarum Gottingensis*, Vol. II, Abraham Vandenhoeck, Göttingen, pp. 159–82.
- Mulholland, J.D. (1969) Numerical studies of lunar motion. *Nature*, **223**, 247–9.
- Mulholland, J.D. (1972) Numerical isolation of flaws in the lunar theory. *Celest. Mech.*, **6**, 242–6.
- Neugebauer, O. (1969) *The Exact Sciences in Antiquity*, Dover Publications Inc., New York.
- Newcomb, S. (1876) *Investigation of Corrections to Hansen's Tables of the Moon, with Tables for Their Application*, U.S. Government Printing Office.
- Newcomb, S. (1895) *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*, Government Printing Office.
- Newcomb, S. (1898) Tables of the motion of the earth on its axis and around the sun, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. VI, No. 1, Bureau of Equipment, Navy Department, Washington.
- de Pontecoulant, G. (1840) *Traité élémentaire de physique céleste, ou précis d'astronomie théorique et pratique, servant d'introduction à l'étude de cette science*, Carillan-Goëury, Paris.
- Ross, F.E. (1917) New elements of mars and tables for correcting the heliocentric positions derived from astronomical Papers, Vol VI, Part IV, in *Astronomical Papers of the American Ephemeris and Nautical Almanac*, Vol. IX, U.S. Government Printing Office, Washington.
- Russell, J.L. (1964) Kepler's laws of planetary motion: 1609–1666. *Brit. J. Hist. Sci.*, **2**, 1–24.
- Saliba, G. (2002) Greek astronomy and the medieval Arabic tradition. *Am. Sci.*, **90**, 360.
- Seidelmann, P.K. (1976) Celestial mechanics, in *Encyclopedia of Computer Science and Technology*, Vol. 4 (eds J. Belzer, A.G. Holzman and A. Kent), Marcel Dekker Inc., New York and Basel.
- Sharma, M.L. (1977) Āryabhata's contribution to Indian astronomy. *Indian J. Hist. Sci.*, **12**, 90.
- Standish, E.M. and Williams, J.G. (2009) Orbital ephemerides of the sun, moon, and planets, in *Explanatory Supplement to the Astronomical Almanac* (eds P.K. Seidelmann, J. A. Bangert and S. Urban), University Science Books, Mill Valley, CA.
- Supplement to the A.E. 1968*, HMNAO, Royal Greenwich Observatory and NAO, U.S. Naval Observatory, Sussex and Washington DC, 1966.
- Szebehely, V.G. and Mark, H. (1998) *Adventures in Celestial Mechanics*, John Wiley & Sons, Inc., New York.
- The Observatory (eds T. Lewis and H.P. Hollis), XXI, Taylor and Francis, London, 1898.
- Turyshchev, S.G., Williams, J.G., Nordvedt, K. Jr. et al. (2004) Years of testing relativistic gravity: where do we go from here? *Lect. Notes Phys.*, **648**, 311–30.
- Will, C.M. (1974) *Experimental Gravitation* (ed. B. Bertotti), Academic Press, New York.

4

Variable Earth Rotation

4.1

Pre 19th Century

The concept of the Earth's rotation was suggested as long ago as the Greek Pythagorean philosophers. The original idea is generally credited to the philosophers of the fourth century BC, Hiketas and Ekphantus of Syracuse, but they may have been influenced by the earlier world system of Philolaus. Aristarchus and his follower, Seleukus, about a century later, also suggested that the Earth rotated on its axis (Dreyer, 1953). That the Earth rotated was not generally believed until the 15th century AD. If the Earth rotated, people and things would be expected to fly off. The general concept was that the Earth was stationary and fixed, and the Sun, Moon, planets, and stars were a limited distance away and moving around the Earth.

Although some individuals in the Middle Ages accepted the Earth's motion, it was Copernicus, Galileo, and Newton that promoted the concepts that the Earth moved around the Sun, that the distances to the planets and stars were large, and that the Earth rotated daily on its axis. Newton's gravitation answered the questions of what held people and things onto the Earth, and provided the basis for computing the motions of celestial bodies, but it was not until 1851 that Léon Foucault provided the first observational 'proof' of the Earth's rotation with his famous pendulum (Tobin, 2003).

The constancy of the Earth's rotation rate was generally not questioned by the 16th century. Chapter 3 outlines the role of celestial mechanics that led to the suggestion of the possible variation in the Earth's rotation. There we see that Flamsteed demonstrated its invariability, at least within his observational accuracy. However, Edmond Halley (1695) used Flamsteed's observations to conclude that the Moon's motion seemed to be speeding up, a fact that could not be explained totally by gravitational forces. Newton suggested that the apparent acceleration of the Moon might be due to changes in the Earth's rotation, but this idea was based on a mistaken idea of interplanetary vapors (Newton, 1713). Immanuel Kant (1754) suggested that the action of the Earth's tides should slow the Earth's rotation rate. Pierre-Simon Laplace (1749–1827), however, determined that the observed deceleration of the Moon in its orbit could be explained gravitationally and that it was

unnecessary to relate this effect to a change in the Earth's rotation. William Ferrel (1853) showed that Laplace had neglected second order effects, and that when they were included a problem remained in explaining the observed lunar motion completely by gravitational forces.

4.2

Secular Variation

The mean longitude of the Moon can be expressed mathematically by the series

$$L = L_0 + n_1 T + n_2 T^2 + \dots, \quad (4.1)$$

where T is time reckoned from some reference epoch.

This expression provides the direction of the Moon as measured along the ecliptic from the vernal equinox. The celestial mechanics explanation for the n_2 term in that expression was provided by Laplace (1786), who tried to explain the entire Moon's motion as being due to planetary perturbations. He pointed out that the disturbing force on the Moon depends on the mean distance of the Earth from the Sun and the eccentricity of the Earth's orbit. Knowing that the eccentricity of the Earth's orbit had been decreasing, Laplace was able to provide a numerical value for n_2 . Adams (1853) pointed out that the tangential component of the disturbing force would also affect the value on n_2 . Even with Adams correction, however, it appeared that the value determined from celestial mechanics was only about half of the value that was derived from ancient eclipse observations.

Although problems were recognized in explaining the secular acceleration of the Moon's motion, it was not until the mid-19th century that the secular retardation of the rate of rotation of the Earth was suggested independently by Charles-Eugene Delaunay (1859, 1866), and William Ferrel (1864) to explain that the secular acceleration of the mean motion of the Moon, determined from gravitational perturbations, was only about half of the observed acceleration determined independently by Dunthorne (1747, 1749), Mayer (1753), and Lalande (1792) from observations over the previous 2500 years. The excess observed secular acceleration was ascribed to tidal retardation of the rotation, which is accompanied by variation in the orbital velocity of the Moon, according to the conservation of momentum.

Tidal deceleration is illustrated in Figure 4.1. The gravitational attraction of the Moon raises an ocean tidal bulge on the Earth. However, the Earth's rotation carries the bulge beyond the line connecting the centers of mass of the two bodies. The gravitational attraction of the Moon on this bulge provides a braking action on the Earth, and we know this as tidal deceleration. The conservation of angular momentum in the Earth-Moon system then causes the moon to be accelerated in its motion around the Earth.

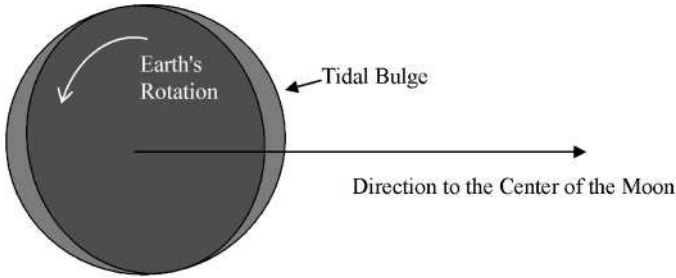


Figure 4.1 Tidal deceleration of the Earth.

Clemence (1971) makes the statement:

The earliest reference that I have found to possible changes in the speed of rotation of the Earth is by Thomson and Tait (1890). They state that Delaunay, who himself made a complete analytical solution of the main problem of the lunar theory (that is, the solution of the three-body problem, in which the bodies are the Earth, the Sun and the Moon), suggested about 1866 that the rotation of the Earth is retarded by tidal friction. They go on to say

‘The ultimate standard of accurate chronometry must (if the human race live on the earth for a few million years) be founded on the physical properties of some body or more constant character than the earth: for instance, a carefully arranged metallic spring ...’ Jacques Lévy, Paris Observatory, based on his personal knowledge of their work, attributes the statement to Thomson (Lord Kelvin).

Simon Newcomb credits Ferrel with being ‘... the first to publish a correct theory of the retardation produced in the rotation of the Earth by the action of the tides and the consequent slow lengthening of the day’ Newcomb (1903a). Brush (1986) points out that ‘Ferrel, strictly speaking, deserves priority for this proposal since he made it at a meeting of the American Academy of Arts and Sciences in Boston in December 1864, a few weeks before Delaunay read his paper to the Académie des Sciences in Paris, but it was Delaunay’s reputation that persuaded astronomers to adopt it.’

Over the years the understanding of the forces involved and the values for the tidal acceleration have improved significantly. Sample values of the kinematical acceleration (d^2L/dt^2), which is twice the value of n_2 , in arcseconds per century squared are given in Table 4.1.

The value compatible with the recent JPL development ephemeris DE 405 of 1997 is $-25.84''/\text{century}^2$. The most recent value for DE421 of 2008 is $-25.85''/\text{century}^2$ (Williams *et al.*, 2008). A more extensive discussion of tidal perturbations

Table 4.1 Sample values of the kinematical acceleration in arcseconds per century².

Reference	Value	Observations
Spencer Jones (1939)	−22	Lunar occultations
Oesterwinter and Cohen (1975)	−38	Lunar occultations
Morrison and Ward (1975)	−26	Lunar occultations
Muller (1976)	−30	Eclipses
Calame and Mulholland (1978)	−24.6	Lunar laser ranging
Ferrari <i>et al.</i> (1980)	−23.8	Lunar laser ranging
Newhall <i>et al.</i> (1988)	−24.90	Lunar laser ranging
Chapront <i>et al.</i> (2002)	−25.826	Lunar laser ranging

and determinations of the values is given in Chapront *et al.*, (2002). From an analysis of the effects on the orbits of near-Earth satellites and the conservation of angular momentum in the Earth-Moon system, Christodoulidis *et al.* (1988) provide the following empirical relation between the tidal acceleration of the Moon and the luni-solar tidal deceleration of the Earth's rotation

$$\dot{\Omega} = (49 \pm 3) \times 0.004869 \dot{n} \times 10^{-22} \text{ rad s}^{-2}.$$

4.3

Irregular Variations in the Earth's Rotation

In the latter part of the nineteenth century Simon Newcomb and Friedrich Ginzel independently investigated observational evidence for irregular variations in the Moon's mean motion and the possibility that these could be explained by variation of the Earth's rotational speed (Stephenson, 2003). While Ginzel investigated records of eclipses (1899), Newcomb established a qualitative correlation between the differences between observed and computed positions of the Moon, and Mercury transit observation residuals. Only about one third of the Moon's residuals could be attributed to variations in the Earth's rotation (Newcomb, 1896). By 1903 Newcomb concluded that the variations in the motion of the moon were really due to causes that had eluded investigation (Newcomb, 1903b). He thought the errors in periodic terms in the theory of the Moon were more likely to explain the differences than variations in the rotation of the Earth. As the Carters point out in their book on Simon Newcomb (2006), in the first chapter of his book *Sidelights on Astronomy*, with editions from 1882 to 1906, Newcomb discussed the unsolved problems of astronomy (Newcomb, 1906). The last problems concern the 'deviations in the movements which astronomers cannot always explain, and which may be due to some hidden causes that, when brought to light, shall lead to conclusions of the greatest importance to our race.' The first deviation he discussed was the rotation of the Earth. Here he said that 'Sometimes for several

years at a time it seems to revolve a little faster, and then again a little slower. The changes are very slight; they can be detected only by the most laborious and refined methods; yet they must have a cause, and we should like to know what that cause is. The moon shows a similar irregularity of motion. For half a century, perhaps through a whole century, she will go around the earth a little ahead of her regular rate, and then for another half century or more she will fall behind. The changes are very small; they would never have been seen with the unaided eye, yet they exist. What is their cause? Mathematicians have vainly spent years of study in trying to answer the question.'

It was not possible to establish the variability of the rotation of the Earth only from observations of the Moon. To do so required comparing the differences between the ephemerides and observations of more than one body. The concept was that irregularities in the Earth's rotation would produce irregularities in the time scales used to make solar system observations that were used to produce the ephemerides. These irregularities would then appear in the ephemerides of all of the solar system objects. However, the variations would be in proportion to their respective mean motions. The accuracies of the observations made it desirable to consider the fastest moving bodies for this investigation. E. W. Brown, in an address to the British Association in Australia in 1913, showed diagrams of the deviations in the longitude of the Moon from its theoretical orbit, along with similar curves for Mercury and the Sun, (Figure 4.2) (Brown, 1914). However, Brown tried to attribute the variations to magnetic forces, and never mentioned the possibility of variations in the rotation of the Earth. Because of the then known connection between sunspot frequency and prevalence of magnetic disturbances, Wolf's sunspot numbers were plotted at the bottom of his diagram. Glauert (1915a, 1915b) further investigated correlations among the Moon, Sun, Mercury, and Venus residuals based on Greenwich observations from 1866 to 1914, but, although his results showed 'satisfactory agreement,' he felt that further investigation was required.

There were a number of difficulties in carrying out these investigations. Observations were made using different techniques at different observatories causing systematic errors among series of observations; they were subject to various accidental errors that varied with time; errors in the positions of the Sun led to errors in the positions of the planets; observations of Mercury and Venus taken before and after inferior conjunction differed systematically; there were uncertainties concerning the location of the equinox, the origin of calculated and observed positions, for the observations and the tables. Meridian observations of Mercury were of limited value, so transit observations (Innes, 1925a) had to be used. Observations of Venus were more numerous and more accurate due to Venus' closer proximity to the Earth. In 1917 Frank Ross and Simon Newcomb (1917) re-discussed the observations of Mars, but found systematic errors in the observations, and because the motion of Mars is slower than that of Venus, the observations were less sensitive to variations in the rotation of the Earth. The basis for the theoretical ephemerides had varied over the years. For example, the solar ephemeris was based on Calini's tables from 1836–1863, Le Verrier's tables from 1864 to

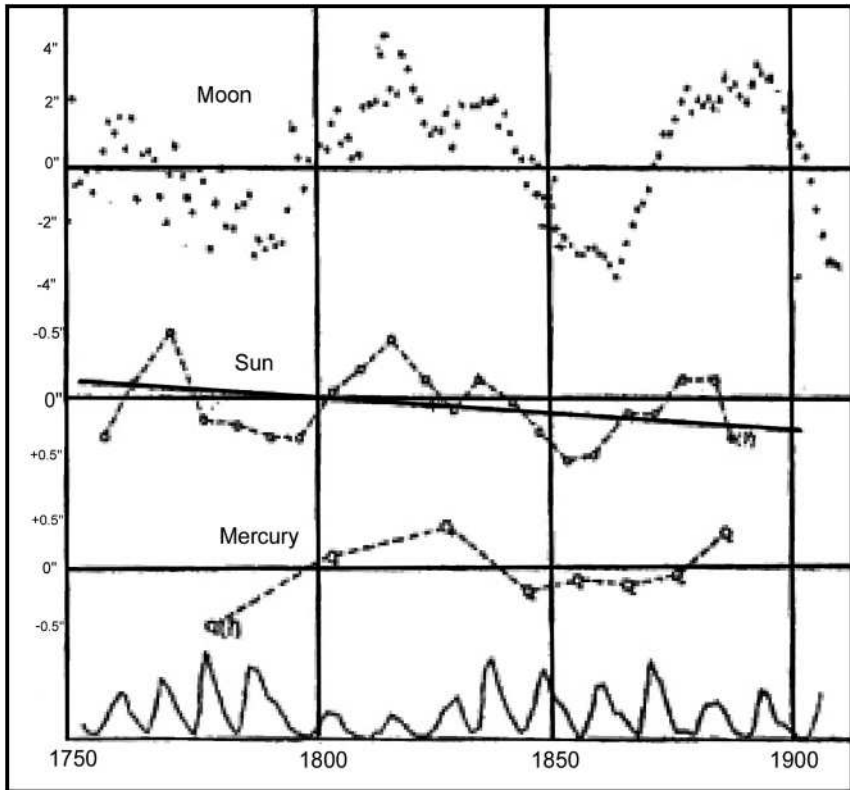


Figure 4.2 Deviations of the longitudes of the Moon, Sun and Mercury from Brown (1914). The bottom curve is Wolf's sunspot numbers that Brown showed because of the known connection between the sunspot frequency and magnetic disturbances.

1900, and Newcomb's tables after 1900. Using these tables led to systematic differences as a function of time.

Innes (1925b) wrote two short notes comparing the time differences indicated by the transits of Mercury, the Moon, eclipses of Jupiter's satellites, and the Sun. He pointed out that, since Einstein had found the explanation for the motion of Mercury's perihelion, all empirical terms ought to be removed and that celestial body motions should be able to be based on gravitational forces only. Empirical terms only caused confusion. He concluded that, while the results were crude, the century 1780 to 1880 was about 30s longer than the century 1800 to 1900. All four sets of data gave qualitatively the same results, indicating variability in the rotation rate of the Earth (Figure 4.3).

Innes (1925a) indicates the need to distinguish between the time observed and the time in a Newtonian sense, which is an absolute, true and mathematical time

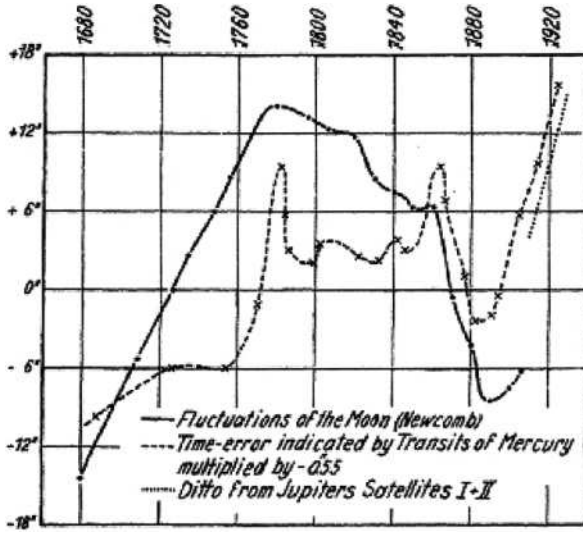


Figure 4.3 Fluctuations in the Moon's motion compared with the time errors determined from transits of Mercury and the motion of Jupiter's satellites from Innes (1925b).

as conceived by Newton as flowing at constant rate, unaffected by the speed or slowness of the motion of material things. He suggested calling rotation time, or observed time, Greenwich Time, and the Newtonian time as World Time or Universal Time, a suggestion that was never adopted. He further suggested that there would be a need to publish the differences as corrections to rotation time, which did become necessary.

Harold Spencer Jones (1926) collected all the data available for a comparison of the residuals of the Sun, Mercury, Venus, and Mars. He then plotted the data, multiplying the planetary residuals by the ratio of the mean motion of the Earth to the mean motion of the planet. If the errors were due to the rotation of the Earth, the four curves should be identical, within the observational errors. Figure 4.4 shows the resulting curves. The curves show a gradual fall from the beginning with a minimum about 1870. Then there is a fairly rapid rise, a slower rise, and a more rapid rise to a maximum at about 1896. Then there is a drop for the rest of the period. The curves are similar qualitatively, and the amplitude changes are about equal. He gave correlation coefficients between the Sun and the planets that range between 0.82 and 0.86, except for the Mercury meridian observations, which are of less accuracy. This indicates that the residuals are from a common cause, namely the variation in the rotation rate of the Earth.

Spencer Jones then prepared a composite curve for the Sun and planets by smoothing the curves and plotting points at two-year intervals. He gave the data weights of 3 for the Sun, 2 for Venus, and 1 for Mercury transits and Mars. That curve is shown in Figure 4.5. The data were compared to longitude residuals of

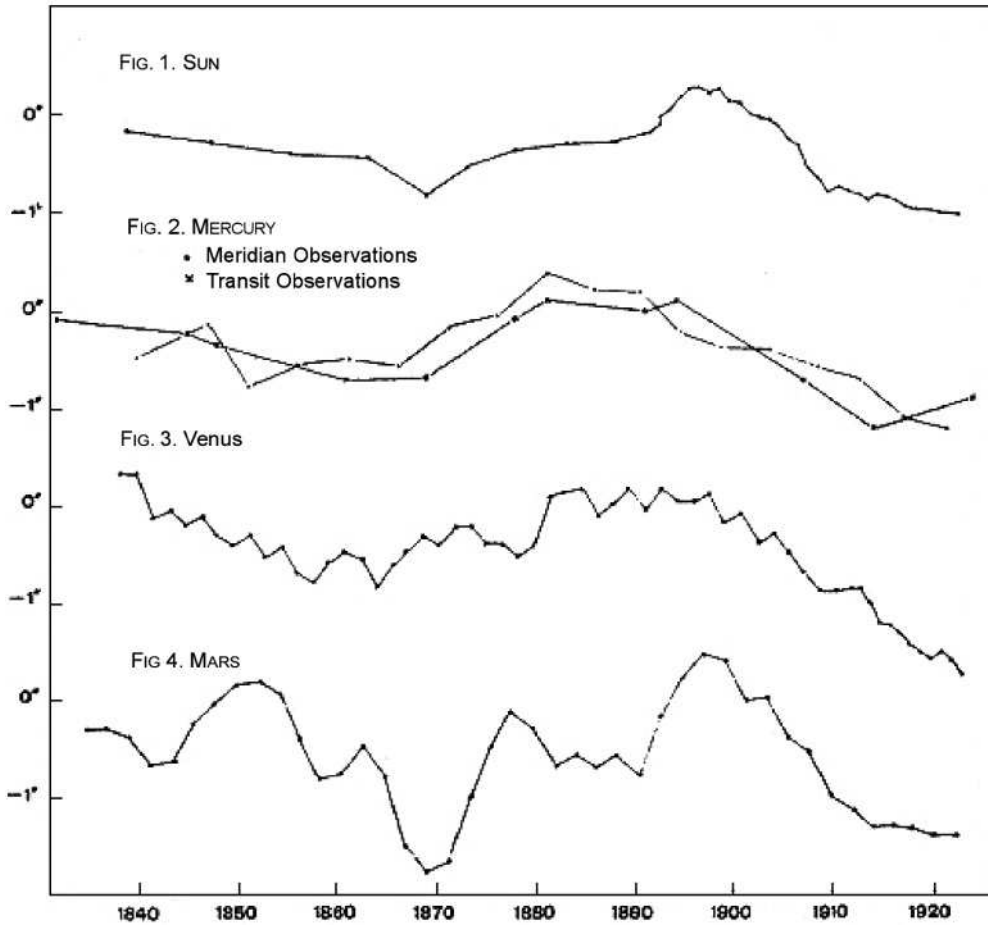


Figure 4.4 Variations in the motion of the Sun and planets from Spencer Jones (1926).

the Moon provided by Dyson and Crommelin (1923). These were residuals from observations made at Greenwich of the Moon's longitude compared to Brown's longitude as modified by Fotheringham (1920), who had corrected the theoretical secular acceleration of Brown to fit the ancient eclipse observations. Fotheringham had introduced a term of period 257 years and semi-amplitude of 13.6 arcseconds. In addition, there were minor fluctuations of irregular period remaining for the Moon. These are shown in Figure 4.5 as the Moon residuals. The scale is adjusted to be approximately the same as the composite curve for the Sun and planets. The similarity of the two curves is evident, but it should be emphasized that, if the great empirical term, introduced by Brown to explain an apparent periodic variation of 257 years, were not removed from the residuals, the curves would not be

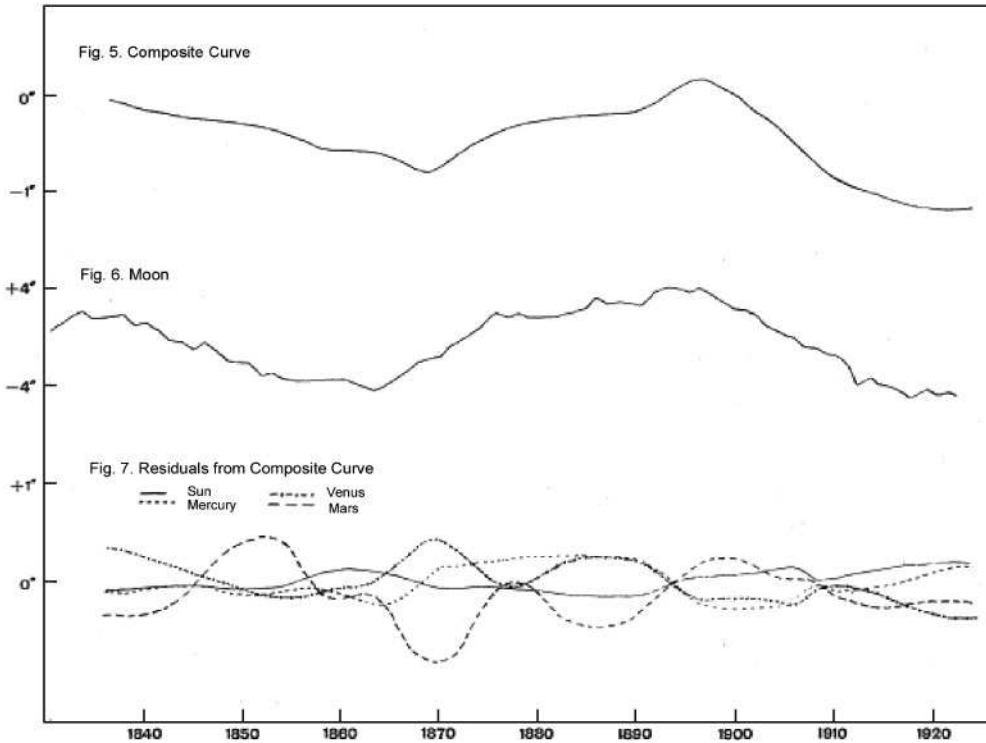


Figure 4.5 Comparison of variations in the Moon's position with the composite planetary variations from Spencer Jones (1926).

similar. If the changes in the longitude residuals between 1897 and 1924 are considered, the ratios of the changes are in reasonable agreement with the theoretical ratios of the mean motions for the planets and the Sun. Spencer Jones concludes that during the period 1836 to 1924 the longitude fluctuations of the Sun, Mercury, Venus, and Mars, and the minor fluctuations of the Moon's longitude, within limits of observational errors, can be attributed to changes in the rate of rotation of the Earth.

de Sitter (1927) addressed two questions concerning the longitude residuals: (i) Are the fluctuations of the longitudes of the Sun and planets equal to those of the Moon in the exact ratios of the mean motions, or diminished by some factor? and (ii) Do the fluctuations of the Sun and planets agree with the total fluctuations of the Moon, or only with minor fluctuations after removing the great empirical term? From this investigation of observational data he found that the secular acceleration of the Sun and planets could be explained by retardation of the Earth's rotation due to tidal friction, but would require a tidal deceleration twice that found from theoretical considerations of the time (Jeffreys, 1924).

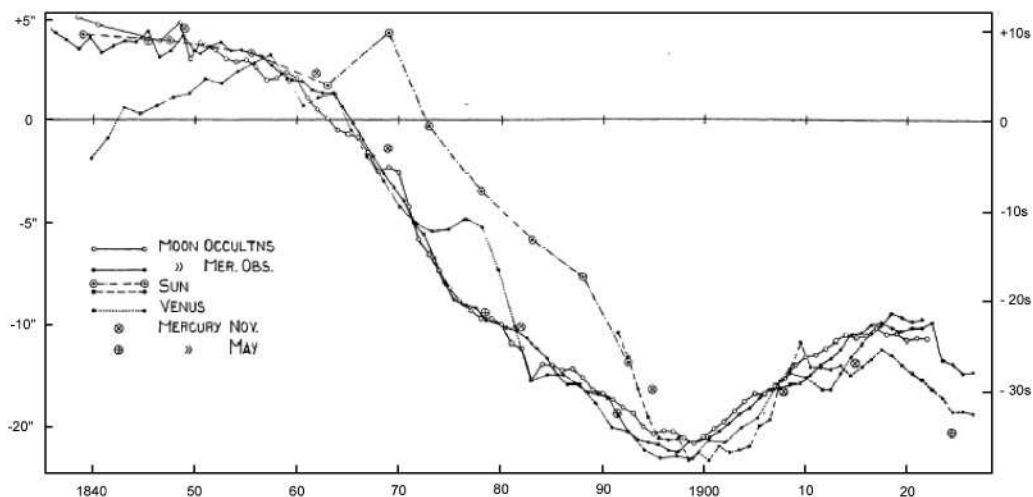


Figure 4.6 de Sitter (1927) plots of observational residuals; left scale is Moon's longitude and right scale is corrections to time.

Considering the great empirical term, the observations of the Sun and Venus were limited to 1835–1925, so short a time that the long-period sinusoidal curve differs little from a straight line. So representations of differences in those cases could be absorbed in solutions for orbital elements. Thus, the only independent observations, besides those of the Moon, were the observations of the transits of Mercury across the Sun's disk. De Sitter concluded that the separation of the great empirical term from the minor fluctuations of the Moon's longitude was artificial. The fluctuations could be explained better by a series of straight lines. He found corrections to Brown's lunar tables, after removing the empirical sine term, and to Newcomb's Tables of the Sun, Mercury and Venus. The combined observational evidence is shown in Figure 4.6. The scale on the left of the figure refers to the Moon's longitude and the scale on the right gives the corrections to the time, corresponding to corrections to the Moon's longitude multiplied by 1.25.

Spencer Jones (1932) discussed the observations of occultations of stars by the Moon from 1672 through 1908 and concluded that the fluctuations in the Moon's longitude could be interpreted as being due to variations in the rate of rotation of the Earth. In 1939 he (Spencer Jones, 1939) established clearly that the fluctuations in the mean longitudes of the Sun, Mercury, and Venus corresponded to the fluctuations in time required to account for the fluctuations in the Moon's mean longitude (see Figures 4.7 and 4.8). Clemence (1943) adopted Spencer Jones's fluctuations in the mean longitudes of Mercury in his discussion of observations of Mercury, and found that the observations agreed with the adopted fluctuations.

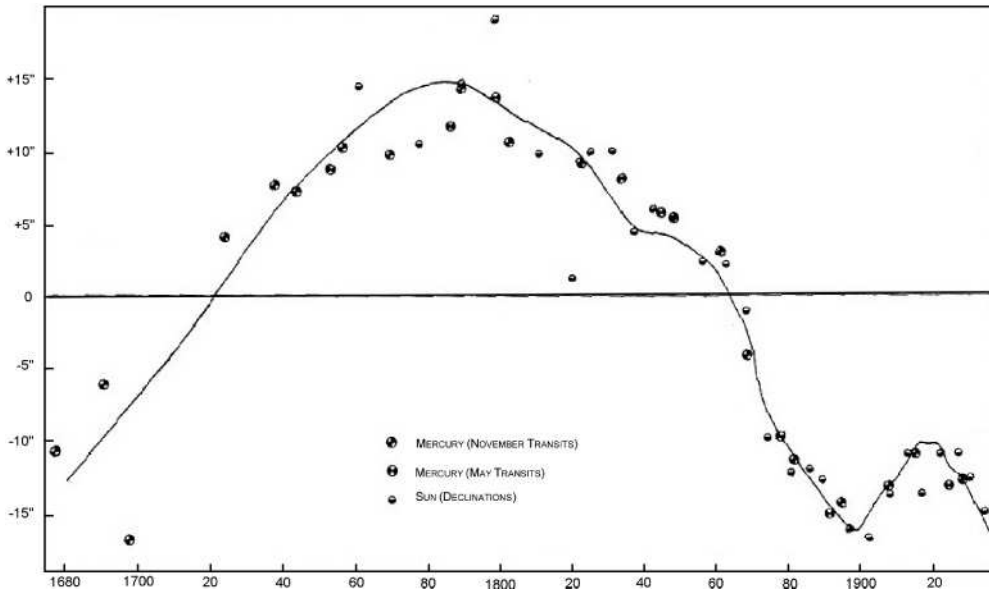


Figure 4.7 Fluctuations in the longitude of the Moon from 1680 from observations of the Moon, Sun, and Mercury from Spencer Jones (1932).

In the 1930s German scientists used newly developed quartz clocks to demonstrate an apparent annual variation in the Earth's rotational speed (Pavel and Uhink, 1935; Scheibe and Adelsberger, 1936). In France Stoyko (1937) tabulated differences between time from pendulum clocks and the rotation of the Earth for three years in the period 1934–1937, and also found an annual variation in the rotation rate of the Earth. He determined that the length of day in January exceeded that in July by 2 milliseconds(ms)/day. By 1950 quartz crystal clocks were used to determine double amplitudes of annual variations of the length of the day of 2.6, 1.8 and 2.8 ms/day by Scheibe and Adelsberger (1950), Finch (1950) and Stoyko (1950) respectively. However, these values were actually too large by factors of two or three. Systematic errors in the assumed positions of stars used to determine time were cause for concern. Smith and Tucker (1953) determined that the annual variation in length of day was less than ± 0.5 ms/day using improved star positions. Markowitz (1955) reported the discovery of fortnightly and monthly variations in the rotational speed, which, using the theoretical calculations of Mintz and Munk (1953), led him to assert that the causes of these variations were meteorological for the annual term and tidal for the others.

Reviews of the Earth's variable rotation rate can be found in Munk and MacDonald (1960), Munk (1966), Lambeck (1980), and the *Explanatory Supplements* (1961, 1992).

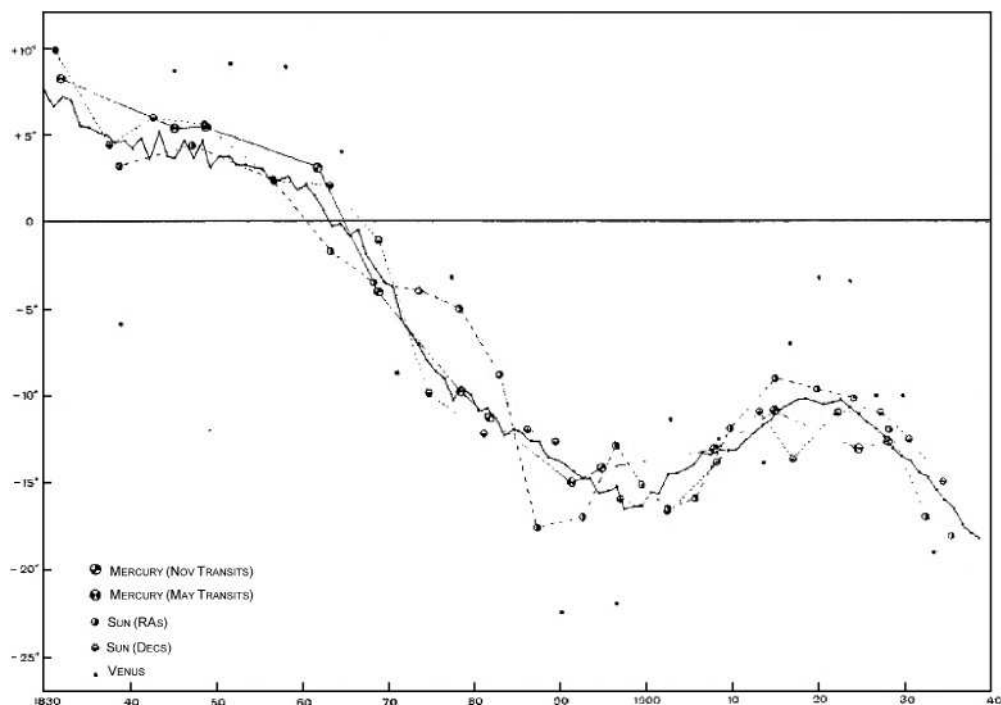


Figure 4.8 Fluctuations in the longitude of the Moon after 1830 from observations of the Moon, Sun, Mercury, and Venus from Spencer Jones (1932).

4.4

Early Explanations for the Variable Rotation

In the face of mounting evidence that the rotational speed of the Earth does change, questions about the physical causes arose. Angular momentum had to be conserved. Computations of the amount of energy required, and/or mass displacements, required to cause fluctuations in the Earth's rotation seemed large and it was hard to explain how such changes could take place without geologists noticing other changes. The astronomical observations became major contributors to learning about the geology and dynamics of the Earth.

While Kant had suggested that the Earth's rotation might be affected by tidal friction, Julius Robert Mayer (1848) described the consequent reaction on the lunar orbit (Brosche, 1984). At the same time that observational evidence for the variable rotational speed was being pursued, George Howard Darwin, the son of the famous biologist was beginning his work on tides and their possible relation to the Earth's rotation (Darwin, 1877, 1879, 1880, 1898). His work built on the ideas of Lord Kelvin (Thomson, 1863). Darwin's work was more concerned with the solid Earth tides, but he did deal with ocean tides. Harold Jeffreys (1924) investigated the effect of ocean tides extensively, providing quantitative estimates of the effect on the Earth's rotation.

In his analysis of the motions of the Moon, Sun, Venus and Mercury, de Sitter (1927) accepted the hypothesis that fluctuations in their longitudes were due to variations in the Earth's rotation and proposed that these variations were due to the combination of two causes. One is a series of abrupt changes in the rate of rotation of the Earth caused by changes in the moment of inertia, perhaps due to expansions and contractions of the Earth. The other cause was proposed to be variability of the coefficient of tidal friction. As Spencer Jones (1939) pointed out in a re-discussion of observations of the Earth's rotation, changes in the moment of inertia of the Earth can increase or decrease its rotation rate and consequently affect the observed longitudes of the Sun, Moon, and planets in proportion to their mean motions. Tidal friction only retards the rotation rate. However, the change in angular momentum of the Earth is compensated by a corresponding change in the angular momentum of the Moon's orbital rotation. The effects on the Moon and planets will be proportional to the mean motions. However, the effect on the longitude of the Moon, being related to the shape of the Moon's orbit, cannot be predicted by theory. Thus, the secular acceleration of the Moon cannot be predicted theoretically.

The apparent irregularities in the Earth's rotation posed a significant problem for the early investigations. Brouwer (1952) discussed the hypothesis that the fluctuations in the Earth's rotation were due to cumulative random changes in its angular velocity. In that article, he mentioned that he and T.E. Sterne had discussed this possibility in about 1935, but concluded that the evidence strongly favored large abrupt changes. He also mentioned that Spencer Jones had speculated that the Earth may be like a pendulum in its behavior and that its rate of rotation is liable to frequent and small irregular changes. However, Brouwer does go on to discuss the explanation that the irregularities are due to apparent variations in the moment of inertia about the Earth's rotation axis. Munk and Revelle (1952) advanced the explanation that the irregularities in the rate of rotation of the Earth were due to electro-magnetic coupling of the mantle to a turbulent core. At about the same time S.K. Runcorn (1954) came to the same interpretation. Vestine (1953) discussed the possibility that these irregularities might be related to motions in the Earth's magnetic field.

Van Den Dungen, Cox and van Mieghem (1949) studied the variation in the moment of inertia of the atmosphere about the polar axis for periodic variations. The preliminary theoretical results were able to account for about a fifth of the observed effects. Munk and Miller (1950) did further studies and found about 15% of the observed effects. Mintz and Munk (1953), however, found that the Earth's reaction would be in reasonable agreement.

4.5

Current Understanding of the Earth's Variable Rotation

We now know that the secular deceleration of the Earth results principally from two causes. Lunar tides should increase the observed length of day by $+2.3 \text{ ms day}^{-1} \text{ cy}^{-1}$. A second effect, caused by the post glacial rebound of portions of the

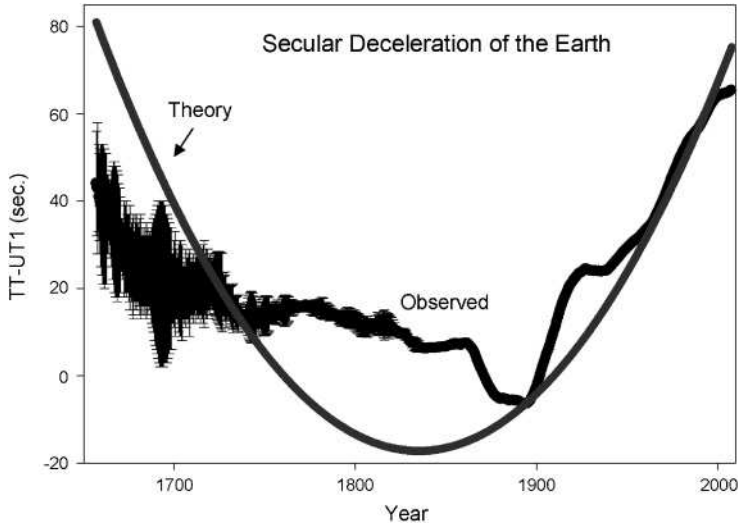


Figure 4.9 Observed and theoretical values of TT–UT1 as a function of time.

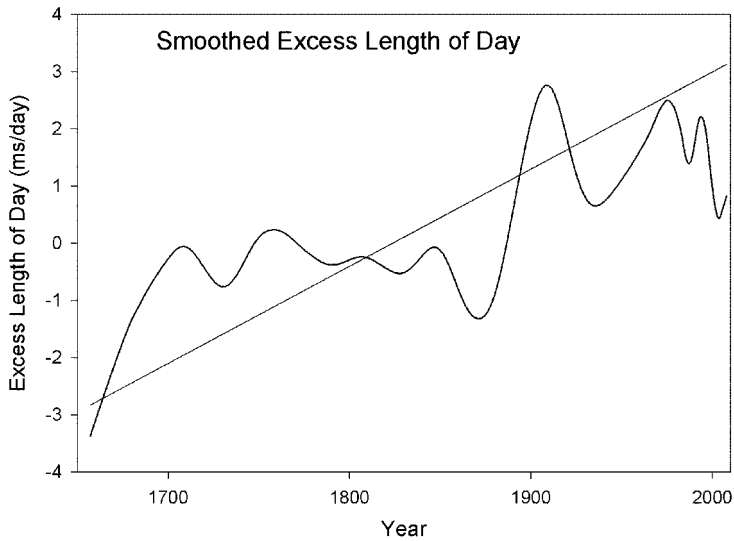


Figure 4.10 Observed and theoretical values of the excess length of day as a function of time.

Earth's crust, actually shortens the length of day by $(-0.6 \pm 0.1) \text{ ms day}^{-1} \text{ cy}^{-1}$. The combination of these two phenomena results in a change in the length of day of $+1.7 \text{ ms day}^{-1} \text{ cy}^{-1}$ (Stephenson and Morrison, 1995). Figure 4.9 shows the difference between UT1 and Terrestrial Time (TT) as a function of time along with the theoretical curve that corresponds to the combination of these two

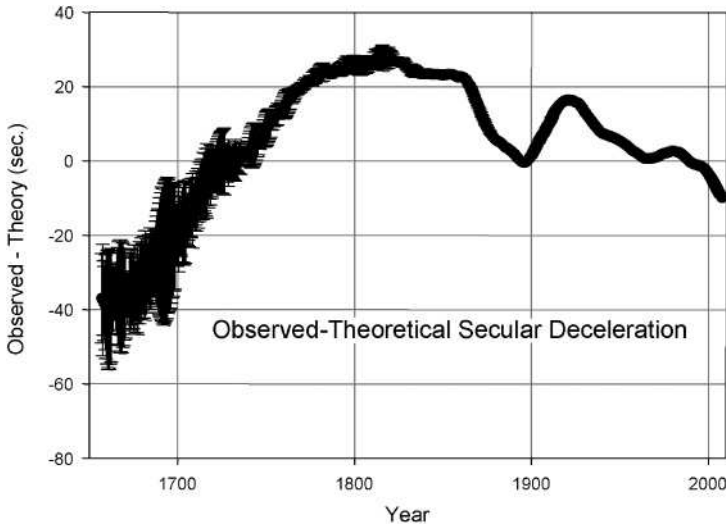


Figure 4.11 Low-frequency irregularities in the TT–UT1.

physical effects. Figure 4.10 shows the expected change in the length of day that corresponds to this combination along with smoothed length of day observations.

Both of these plots show that there are significant departures from a purely secular deceleration of the Earth's rotation. The decadal irregularities, as shown in Figure 4.11, are attributed to the interactions of the Earth's mantle and the liquid core.

This is based on observations of the magnetic field at the Earth's surface along with theoretical analyses of motions in the Earth's core (Dehant, de Viron and van Hoolst, 2005; Mound and Buffett, 2003).

We now know that the Earth is subject to variations at many frequencies from decadal to sub-daily, and that these variations have many geophysical and meteorological causes. These were outlined by Eubanks (1993), and Figure 4.12, taken in part from that review, shows schematically the various components of the Earth's variable rotation. It shows that the annual and higher frequency variations in the Earth's rotation are closely correlated with the total atmospheric angular momentum (AAM). Other contributors include ocean angular momentum and hydrology. Periodic variations due to tides are also present.

4.6 Consequences

In the early 20th century it became apparent that 'astronomical time,' based on the Earth's rotation, and used for practical astronomical computations, differed from a 'uniform' or 'Newtonian' time, which is the independent variable for cele-

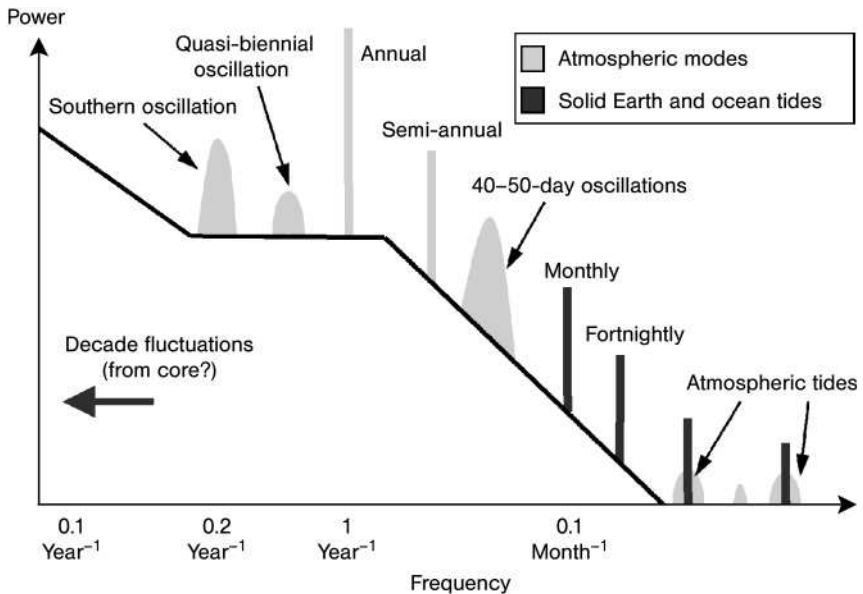


Figure 4.12 Schematic representation of higher-frequency variations in the Earth's rotation.

tial mechanics. Recognizing this fact, de Sitter went on to provide observationally determined corrections to astronomical time to obtain a uniform time. These included both a secular term and irregular corrections given in a table covering the period 1640–1926.5.

With the recognition that the Earth's rotation was not uniform, the method of rating clocks using star observations was limited in accuracy by the variability of the Earth's rotation. So as clocks became more accurate than the Earth, ensembles of clocks were necessary to establish accurate time. Crystal-controlled clocks could smooth out the random observational errors from night to night. The crystal oscillators that were used as time standards varied in frequency from day to day by only a few parts in 10^{10} . To determine time based on the rotation of the Earth, it was necessary to observe selected stars to obtain the hour angle of the equinox based on a catalog of the directions of the stars. So the observations of the positions of stars in their diurnal circuits determined apparent sidereal time referred to observers' instantaneous local meridians. Different instruments were used for these observations, such as meridian or transit circles, photographic zenith tubes, and Danjon astrolabes at optical wavelengths. Optical observations could determine time to a few milliseconds of time.

The recognition of the variability of the rotation of the Earth also led to some far-reaching consequences. It meant that measuring time based on the Earth's rotation using angle measurements of celestial objects would no longer be suitable for those who required the most accurate time scales. It also meant that an observing program would be necessary to monitor the Earth's rotation (Chapter 5) separately from a uniform time based on mechanical

clocks for those who required knowledge of the Earth's orientation in space. So from here on we must separate the concept of time and the determination of time scales into the different types, and trace the developments in the separate cases.

To provide accurate directions to solar system objects, astronomers developed a uniform time for ephemerides and celestial mechanics. This new time scale, defined in terms of the Sun's orbital motion itself, was given the name 'Ephemeris Time' (Chapter 6). For everyday practical use, however, time was provided by clocks. The most accurate mechanical clocks were superseded by the introduction of quartz crystal clocks and eventually atomic clocks, providing accuracy improvements by orders of magnitude. Practical timekeeping would no longer be the job of astronomers alone. The familiar connection of time to the Sun continued to require astronomical measurements to 'regulate' the more uniform atomic time, but precise time measurements would now be more and more in the hands of the physicists and engineers.

On the other hand, the development of more accurate reference frames, to accommodate the improving solar system ephemerides and the introduction of artificial Earth satellites and interplanetary space travel, made it even more necessary to make observations to determine the rotation of the Earth as one of the components of our knowledge of the Earth's orientation in space. Mathematical models alone could not provide information with the accuracy needed for practical use in these rapidly developing areas. Astronomical observations were required to close the gap between models and observations. This requirement led to the development of new observational techniques to provide improved measurements, which led, in turn, to the development of areas of geophysics to improve our knowledge of Earth systems, which was made possible by the improved observations. The field of Earth orientation sciences bridging astronomy, geodesy and geophysics came into existence.

References

- Adams, J.C. (1853) On the secular variation of the moon's mean motion. *Philos. Trans. R. Soc. London*, CXLIII, 397–406.
- Brosche, P. (1984) Tidal friction in the Earth-Moon system. *Philos. Trans. R. Soc. London A*, 313, 71–5.
- Brouwer, D. (1952) A study of the changes in the rate of rotation of the earth. *Astron. J.*, 57, 125–46.
- Brown, E.W. (1914) Address on cosmical physics. *Brit. Assoc. Rep. Aust.*, 311–21; *Science, New Series*, 40, 389–401.
- Brush, S.G. (1986) Early history of selenogony, in *Origin of the Moon; Proceedings of the Conference, Kona, HI, October 13–16, 1984*, Lunar and Planetary Institute, Houston, TX, pp. 3–15.
- Calame, O. and Mulholland, J.D. (1978) Lunar Tidal acceleration determined from laser range measures. *Science*, 199, 977–8.
- Carter, B. and Carter, M.S. (2006) *Simon Newcomb, America's Unofficial Astronomer Royal*, Mantanzas Publishing, St. Augustine, FL.
- Chapront, J., Chapront-Touze, M. and Francoou, G. (2002) A new determination of lunar orbital parameters, precession constant, and tidal acceleration from LLR measurements. *Astron. Astrophys.*, 387, 700–9.
- Christodoulidis, D.C., Smith, D.E., Williamson, R.G. and Klosko, S.M. (1988) Observed tidal braking in the Earth/Moon/Sun system. *J. Geophys. Res.*, 93, 6216–36.

- Clemence, G.M. (1943) The motion of mercury 1765–1937, in *Astronomy Papers of the American Ephemeris and Nautical Almanac*, Vol. XI, Part I, Government Printing Office, Washington.
- Clemence, G.M. (1971) The concept of ephemeris time: a case of inadvertent plagiarism. *J. Hist. Astron.*, **2**, 73–9.
- Darwin, G.H. (1877) On the influence of geological changes on the Earth's axis of rotation. *Philos. Trans. R. Soc. London*, **167**, 271.
- Darwin, G.H. (1898) *Tides and Kindred Phenomena in the Solar System*, Houghton, Mifflin and Company, Boston and New York.
- Darwin, G.H. (1879) On the precession of a viscous spheroid and on the remote history of the earth. *Philos. Trans. R. Soc. London*, **170**, 447–530.
- Darwin, G.H. (1880) On the secular change of the orbit of a satellite revolving about a tidally distorted planet. *Philos. Trans. R. Soc. London*, **171**, 713–891.
- Dehant, V., de Viron, O., van Hoolst, T. (2005) Poincaré flow in the Earth's core, in *Journées 2004–systèmes de référence spatio-temporels. Fundamental astronomy: new concepts and models for high accuracy observations, Paris, 20–22 September 2004* (ed. N. Capitaine), Observatoire de Paris, Paris.
- Delaunay, C.-E. (1859) *Comptes Rendus Acad. Sci. Paris* (Séance du 25/4/1859), **48**, 817.
- Delaunay, C.-E. (1866) *Conférence sur l'astronomie et en particulier sur le ralentissement du mouvement de rotation de la terre*, G. Bailliere, Paris.
- Dreyer, J.L.E. (1953) *The History of Astronomy from Thales to Kepler*, Dover Publications.
- Dunthorne, R. (1747) A letter from Mr. Richard Dunthorne, to the Rev. Mr. Cha. Mason, F. R. S. and Woodwardian Professor of Nat. Hist. at Cambridge, concerning the Moon's Motion. *Philos. Trans.*, **44**, 412–20.
- Dunthorne, R. (1749) A letter from the Rev. Mr. Richard Dunthorne to the Reverend Mr. Richard Mason F. R. S. and keeper of the Woodwardian Museum at Cambridge, concerning the Acceleration of the Moon. *Philos. Trans.*, **46**, 162–72.
- Dyson, F. Sir and Crommelin, A.C.D. (1923) The Greenwich observations of the moon. *Mon. Not. R. Astron. Soc.*, **83**, 359–70.
- Eubanks, T.M. (1993) Variations in the Orientation of the earth, in *Contributions of Space Geodesy to Geodynamics: Earth Dynamics: Geodynamic Series*, Vol. 24 (eds D.E. Smith, D.L. Turcotte), American Geophysical Union, Washington, DC, p. 1.
- Explanatory Supplement to the Astronomical Almanac* (ed. P. Kenneth Seidelmann), University Science Books, Mill Valley, CA, 1992.
- Explanatory Supplement to The Astronomical Ephemeris and The American Ephemeris and Nautical Almanac*, Her Majesty's Stationery Office, London, 1961.
- Ferrari, A.J., Sinclair, W.S., Sjogren, W.L., Williams, J.G. and Yoder, C.F. (1980) *J. Geophys. Res.*, **85**, 3939.
- Ferrel, W. (1853) On the effect of the sun and moon upon the rotatory motion of the earth. *Astron. J.*, **3**, 138–41.
- Ferrel, W. (1864) Note on the influence of the tides in causing an apparent secular acceleration of the moon's mean motion. *Proc. Am. Acad. Art. Sci.*, **VI**, 379–83, 390–3.
- Finch, H. (1950) On a periodic fluctuation in the length of the day. *Mon. Not. R. Astron. Soc.*, **110**, 3.
- Fotheringham, J.K. (1920) The longitude of the Moon from 1627 to 1918. *Mon. Not. R. Astron. Soc.*, **80**, 289.
- Ginzel, F.K. (1899) Bemerkungen über den Werth der alten historischen Sonnenfinsterniss für die Mondtheorie. *Astron. Nachr.*, **150**, 1.
- Glauert, H. (1915a) The rotation of the earth. *Mon. Not. R. Astron. Soc.*, **75**, 489–95.
- Glauert, H. (1915b) The rotation of the earth. *Mon. Not. R. Astron. Soc.*, **75**, 685–7.
- Halley, E. (1695) Some account of the ancient state of the city of Palmyra, with short remarks upon the inscriptions found there. *Philos. Trans. R. Soc.*, **19**, 160.
- Innes, R.T.A. (1925a) Transits of mercury 1677–924. *Union Obs. Circ.*, **65**, 303–324.
- Innes, R.T.A. (1925b) Variability of the earth's rotation. *Astron. Nachr.*, **25**, 109.

- Jeffreys, H. (1924) *The Earth: Its Origin, History and Physical Constitution*, Cambridge University Press.
- Kant, I. (1754) 'Untersuchung der Frage, ob die Erde in ihrer Umdrehung um die Achse, wodurch sie die Abwechselung des Tages und der Nacht hervorbringt, einige Veränderung seit den ersten Zeiten ihres Ursprungs erlitten habe und woraus man sich ihrer versichern könne, welche von der Königlichen Akademie der Wissenschaften zu Berlin zum Preise für das jetztlaufende Jahr aufgegeben worden,' English: 'Investigation of the Question, Whether the Axial Rotation of the Earth, through which Day and Night are brought about, has Changed since its Beginning, and How One Can be Certain of this, which the Royal Academy of Sciences in Berlin has offered a Prize for the current year,' in *Wöchentliche Königsbergische Frag- und Anzeigungs-Nachrichten* #23 (June 8) and #24 (June 15). [Ak. 1: 185–91.], Reußner/Hartung, Königsberg.
- Lalande, J. (1792) *Traité d'astronomie* (2 Vols, 1764 enlarged. edition, 4 Vols, 1771–1781; 3rd edn, 3 Vols, 1792), P. Didot.
- Lambeck, K. (1980) *The Earth's Variable Rotation: Geophysical Causes and Consequences*, Cambridge University Press, Cambridge.
- Laplace, P.-S. (1786) Sur l'équation séculaire de la Lune. *Mém. Acad. R. Sci.*, 235.
- Markowitz, W. (1955) The annual variation in the rotation of the Earth, 1951–4. *Astron. J.*, 59, 69.
- Mayer, R. (1848), *Beiträge zur Dynamik des Himmels*, Chapter 8, Landherr, Heilbronn.
- Mayer, T. (1753) *Novae Tabulae motuum Solis et Lunae*, in *Commentarii Societatis Regiae Scientiarum Gottingensis*, Vol. II, Abraham Vandenhoeck, Göttingen, pp. 159–82.
- Mintz, Y. and Munk, W. (1953) The effect of winds and bodily tides on the annual variation in the length of day. *Mon. Not. R. Astron. Soc.*, 113, 789.
- Morrison, L.V. and Ward, C.G. (1975) An analysis of the transits of mercury-1677–1973. *Mon. Not. R. Astron. Soc.*, 173, 183–206.
- Mound, J.E. and Buffett, B.A. (2003) Interannual oscillations in length of day: implications for the structure of the mantle and core. *J. Geophys. Res. Solid Earth*, 108 (B7), ETG2-1. CiteID 2334, DOI: 10.1029/2002JB002054.
- Muller, P.M. (1976) *JPL Spec. Publ.*, 43, 36.
- Munk, W.H. (1966) Variation of the earth's rotation in historical time, *The Earth-Moon System* (eds B.G. Marsden and A.G.W. Cameron), Plenum Press, New York.
- Munk, W.H. and MacDonald, G.J.F. (1960) *The Rotation of the Earth*, Cambridge University Press, Cambridge.
- Munk, W.H. and Miller, R.L. (1950) Variations in the earth's angular velocity resulting from fluctuations in atmospheric and ocean circulation. *Tellus*, 2, 93–101.
- Munk, W.H. and Revelle, R. (1952) On the geophysical interpretation of irregularities in the rotation of the earth. *Mon. Not. R. Astron. Soc. Geophys. Suppl.*, 6, 331.
- Newcomb, S. (1896) *Comptes Rendus Acad. Sci. Paris*, tome 1, cxxii, 1238.
- Newcomb, S. (1903a) *The Reminiscences of an Astronomer*, Houghton Mifflin and Company, Boston, MA and New York.
- Newcomb, S. (1903b) On the desirableness of re-investigation of the problems growing out of the mean motion of the moon. *Mon. Not. R. Astron. Soc.*, 63, 318–24.
- Newcomb, S. (1906) *Side-lights on Astronomy and Kindred Fields of Popular Science. Essays and Addresses*, Harper & Brothers, London and New York.
- Newhall, XX, Williams, J.G. and Dickey, J.O. (1988) Earth rotation from lunar laser ranging, in *The Earth's Rotation and Reference Frame for Geodesy and Geodynamics* (ed. A.K. Babcock and G.A. Wilkins), Kluwer Academic Publishers, Dordrecht, p. 159.
- Newton, I. (1713) *Philosophia Naturalis Principia Mathematica*, 2nd edn, S. Pepys, London, p. 481.
- Oesterwinter, C. and Cohen, C.J. (1975) New orbital elements for moon and planets. *Celest. Mech.*, 5, 317–95.
- Pavel, F. and Uhink, W. (1935) Die Quarzuhren des Geodätischen Instituts in Potsdam. *Astron. Nachr.*, 257, 365–90.
- Ross, F.E. and Newcomb, S. (1917) New elements of Mars and tables for correcting the heliocentric positions derived from Astronomical papers. *Astron Papers of the AENA*, IX, Part II.
- Runcorn, S.K. (1954) The earth's core. *Trans. Am. Geophys. Union*, 35, 49.

- Scheibe, A. and Adelsberger, U. (1936) Nachweis von Schwankungen der astronomischen Tageslänge mittels Quarzuhren. *Phys. Zeitschrift*, **37**, 38.
- Scheibe, A. and Adelsberger, U. (1950) Die Gangleistungen der PTR-Quarzuhren und die jährliche Schwankung der astronomischen Tageslänge. *Zeitschrift für Physik*, **127**, 416.
- de Sitter, W. (1927) On the secular accelerations and the fluctuations of the longitudes of the moon, the sun, Mercury and Venus. *Bull. of Astron. Inst. of Netherlands*, IV, 21–38.
- Smith, H. and Tucker, R. (1953) The annual fluctuation in the rate of rotation of the Earth. *Mon. Not. R. Astron. Soc.*, **113**, 251.
- Spencer Jones, H. (1926) The rotation of the earth. *Mon. Not. R. Astron. Soc.*, **87**, 4–31.
- Spencer Jones, H. (1932) Discussion of observations of occultations of stars by the moon, 1672–1908 being a revision of Newcomb's 'researches on the motion of the moon, part II'. *Ann. Cape Obs.*, XIII, 3–70.
- Spencer Jones, H. (1939) The rotation of the earth and the secular acceleration of the sun, moon, and planets. *Mon. Not. R. Astron. Soc.*, **99**, 541.
- Stephenson, F.R. (2003) Historical eclipses and earth rotation. *Astron. Geophys.*, **44**, 222–7.
- Stephenson, F.R. and Morrison, L.V. (1995) Long-term fluctuations in the earth's rotation: 700 BC to AD 1990. *Philos. Trans. Phys. Sci. Eng.*, **351**, 165–202.
- Stoyko, N. (1937) Sur la périodicité dans l'irrégularité de la rotation de la Terre. *Comptes Rendus Acad. Sci.*, **205**, 79–81.
- Stoyko, N. (1950) Sur la variation saisonnière de la rotation de la terre. *Comptes Rendus Acad. Sci.*, **230**, 514.
- Thomson, W. (1863) On the rigidity of the earth. *Phil. Trans. R. Soc. London*, **153**, 573–82.
- Thomson, W. and Tait, P.G. (1890) *Treatise on Natural Philosophy*, Cambridge University Press, Cambridge, para 405 (footnote).
- Tobin, W. (2003) *The Life and Science of Léon Foucault: The Man Who Proved the Earth Rotates*, Cambridge University Press.
- Van den Dungen, F.H., Cox, F.J. and van Mieghem, J. (1949) *Bull. Acad. Belg. Cl. Sci.*, **35**, 642–55.
- Vestine, E.H. (1953) On variations of the geomagnetic field, fluid motions, and the rate of the earth's rotation. *J. Geophys. Res.*, **58**, 127.
- Williams, J.G., Boggs, D.H. and Folkner, W.M. (2008) DE421 lunar orbit, physical librations, and surface coordinates. IOM 335-JW, DB, WF-20080314-001.

5

Earth Rotation and Polar Motion

5.1

Earth Orientation

As we learned more about the Earth's variable rotation, it became clear that the Earth was becoming less suitable as a means of providing a uniform time scale. Nevertheless, information regarding the Earth's rotation angle and kinematics in space remains a critical element in describing the orientation of the planet in space and relating reference frames on the Earth to those in space.

Terrestrial and celestial reference systems are composed of (i) a specified origin, (ii) the directions of fundamental axes, and (iii) a set of conventional models, procedures, and constants used to realize the system. A reference frame is the realization of that system through a list of coordinates. Earth orientation describes the procedure and models used to relate a terrestrial geodetic reference system to a celestial reference system. The rigorous details are outlined in the publications of the International Earth Rotation and Reference Systems Service (IERS), specifically in IERS Conventions (2003) and its updates, which are available electronically at <http://www.iers.org/iers/products/conv/>.

Figure 5.1 shows conceptually a terrestrial reference frame (red axes) and a celestial reference frame (yellow axes). Celestial reference systems are specified by astronomically defined directions and origins. Most modern systems are considered to have their origins at the barycenter of the solar system, their polar axes being related in some way to the rotational axis of the Earth. The second axis then lies in the equatorial plane perpendicular to the polar axis and is directed toward a fiducial point in that plane, formerly the vernal equinox. The third axis is chosen to complete a right-handed orthogonal system. A list of the positions and possible motions of astronomical objects then comprises the related celestial reference frames.

Historically, the fixed celestial reference frame was defined by the positions of optically bright, nearby stars at some epoch. A moving reference system was based on the solar system dynamics. Thus, the x axis was specified by the vernal equinox, the intersection of the ecliptic and equatorial planes, both of which were in motion. Precession and nutation were described by conventionally adopted models such as the IAU 1976 Astronomical Constants and the IAU 1980 Theory

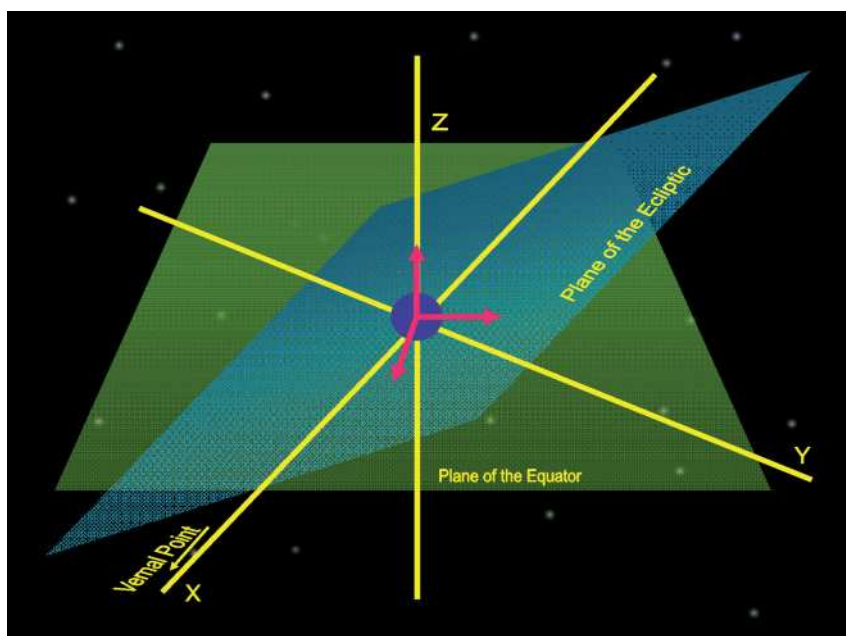


Figure 5.1 Concept of terrestrial and celestial reference frames. The x axis of the celestial frame follows the direction to the vernal equinox and the x axis in the rotating terrestrial frame is in the direction of the zero meridian, near Greenwich.

of Nutation. ‘Mean positions’ of stars for specific epochs were based on the adopted precession. ‘True positions’ of date were determined using the adopted theory of nutation. Terrestrial positions were determined based on the rotation of the Earth and polar motion. Precession and nutation models could be developed using adopted theoretical geophysical models. Earth rotation and polar motion could only be determined from observations and projected forward for limited periods of time.

In 1992 the International Astronomical Union (IAU) designated a standard, fixed, epoch-independent, celestial reference system and frame called the International Celestial Reference System (ICRS) and the International Celestial Reference Frame (ICRF), respectively (Ma and Feissel, 1997). The ICRF is made up of the adopted directions to distant radio sources. Positions of optical stars, consistent with that frame, are provided by the Hipparcos Catalog (Perryman *et al.*, 1997).

Terrestrial reference systems generally have their origins at the center of mass of the Earth with their polar axes related to the direction of an axis fixed with respect to the Earth’s crust. The origin of longitudes in the equatorial plane provides the second direction. In 1884 the Greenwich meridian defined astronomically was chosen to provide this origin, but improvements in geodetic accuracy since then made this definition obsolete. Nevertheless, the origin of longitudes is,

in practice, very near the site identified by the 1884 decision. The third axis is chosen to complete a right-handed orthogonal system. The terrestrial frames are comprised of a list of site coordinates and, in more recent times, their possible motions, which is similar to the use of stellar positions to define the celestial reference system. The International Terrestrial Reference System (ITRS) and the International Terrestrial Reference Frame (ITRF) (Boucher *et al.*, 2004), maintained by the IERS, are accepted as the international standards.

The ICRF is considered to have its origin at the solar system barycenter. Consequently, it is one component of a Barycentric Celestial Reference System (BCRS), and the ICRS is considered to be a specific realization of a BCRS. Observations are generally made in the terrestrial system, whose origin is at the geocenter. In doing this we make use of a Geocentric Celestial Reference System (GCRS) that is kinematically nonrotating with respect to a BCRS, so the orientation of the GCRS is the orientation of the ICRS. The origin of the GCRS moves nonlinearly with respect to the BCRS, but it has fixed directions with respect to the extragalactic sources. Thus, there is a Coriolis-like effect from relativistic theory in the transformations, if referred to the BCRS. This effect is called geodesic precession and nutation and is included in the precession-nutation computation.

The terrestrial system is rotating in the celestial system, and its orientation in that system is affected by precession, nutation, polar motion, and variations in the Earth's rotational speed. Modeling these effects with precision requires the introduction of a number of complications. The fact that the Earth is not strictly a rigid body means that nonrigid body effects need to be included. The fact that the Earth's core experiences a free wobble with respect to the mantle means that existing geophysical models of nutation may not account for all of the observed motions. Further, motions caused by redistribution of mass in the Earth, its oceans and atmosphere, along with relatively high-frequency variations in global meteorology and hydrology, may need to be taken into account.

To facilitate the transformation between a barycentric celestial reference frame and a geocentric terrestrial reference frame, a Celestial Intermediate Reference System (CIRS) is used. It is a geocentric reference system related to the GCRS by a time-dependent rotation taking into account precession-nutation. It is defined by the intermediate equator of the Celestial Intermediate Pole (CIP) and the Celestial Intermediate Origin (CIO) for a specific date. The CIP is a geocentric equatorial pole, whose direction results from (i) the part of the precession-nutation model with periods greater than 2 days, (ii) the retrograde diurnal part of polar motion (including the free core nutation), and (iii) the frame bias. Its ITRS position results from (i) the part of polar motion which is outside the retrograde diurnal band in the ITRS and (ii) the motion in the ITRS corresponding to nutation components with periods less than 2 days. The CIP is a conventionally defined pole separating the motion of the pole of the ITRS in the ICRS into the celestial motion of the CIP (precession/nutation), including all the terms with periods greater than 2 days in the Celestial Reference System (CRS) (frequencies between -0.5 cycles per sidereal day (cpsd) and $+0.5$ cpsd), and the terrestrial motion of the CIP (polar motion), including all the terms outside the retrograde diurnal band in the

Terrestrial Reference System (TRS) (frequencies lower than -1.5 cpsd or greater than -0.5 cpsd). In the IAU 1980 Theory of Nutation the pole was called the Celestial Ephemeris Pole (CEP), but the name was changed to be consistent with new system nomenclature. The actual motion of the CIP is realized by the IAU 2000A precession-nutation model plus time-dependent corrections provided by the International Earth Rotation and Reference System Service (IERS). The division of periodic terms with periods more than and less than 2 days is arbitrary.

The CIO is the origin of right ascensions on the intermediate equator in the CIRS. It is the nonrotating origin in the GCRS that was originally set close to the GCRS meridian and throughout 1900–2100 stays within 0.1 arc seconds of this alignment. The CIO was located on the true (CIP) equator of J2000.0 at a point 2.012 mas from the ICRS prime meridian at right ascension 00h 00m 00s.00013416 in the ICRS. As the true equator moves in space, the path of the CIO in space is such that the point has no instantaneous east-west velocity along the true equator. In contrast, the equinox has instantaneous velocity along the equator.

Similarly, a Terrestrial Intermediate Reference System (TIRS) is used in transforming between frames. This is a geocentric reference system defined by the intermediate equator of the CIP and the Terrestrial Intermediate Origin (TIO). It is related to the International Terrestrial Reference System (ITRS) by polar motion and a small, slowly varying quantity called the TIO locator. It is related to the Celestial Intermediate Reference System (CIRS) by a rotation called the Earth Rotation Angle (ERA) around the CIP, which defines the common z -axis of the two systems. The TIO was originally set at the International Terrestrial Reference Frame (ITRF) origin of longitude, and throughout 1900–2100 stays within 0.1 mas of the ITRF zero-meridian.

With the introduction of the new reference system in the 1992–2004 period, the ICRS replaced the reference system based on the positions and motions of the FK5. The CIO replaced the moving vernal equinox; the TIO replaced the Greenwich Meridian; the Earth Rotation Angle (ERA) replaced the Greenwich Sidereal Time. The alternative system based on the equinox, mean and true positions, and the Greenwich Mean Sidereal Time is still supported and when properly applied can provide equivalent accuracies (Kaplan, 2005).

The transformation between celestial and terrestrial frames is specified by five angles called Earth orientation parameters. Three would be sufficient, but five angles are used in order to describe the physical processes involved and to make the transformations easier to apply. Two angles are used to model the changing direction of the CIP due to the precession and nutation of the Earth. These phenomena are driven by the gravitational attraction of the solar system bodies, principally the Sun and the Moon, on the nonspherical Earth. Precession refers to the aperiodic portion of the motion and nutation refers to the periodic portion. Both motions depend on the positions of the solar system bodies and the internal structure of the Earth, but they can be modeled mathematically with reasonable accuracy.

Two more angles are used to describe the motion of the CIP with respect to the Earth's crust. This phenomenon is called 'polar motion' and is driven by geophy-

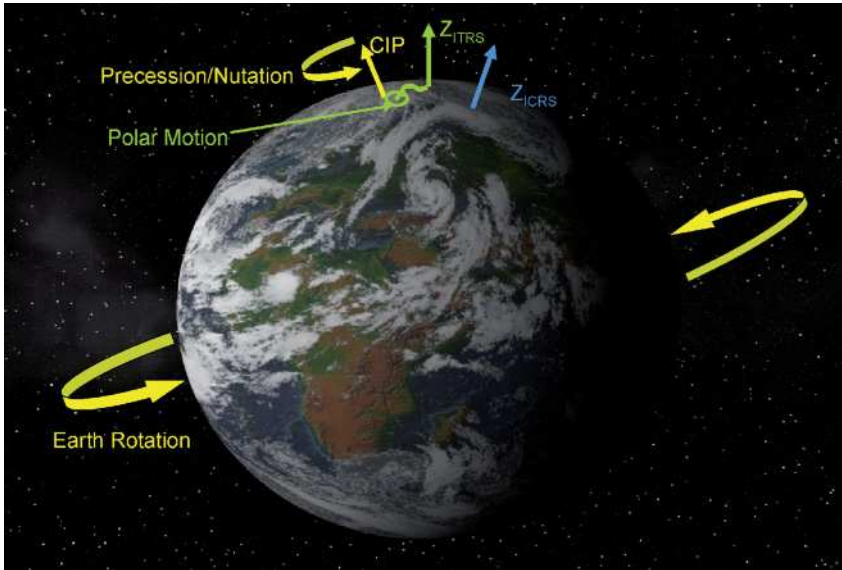


Figure 5.2 Reference system axes and motions related to the Earth's orientation in space.

sical and meteorological variations within the Earth and its atmosphere. Polar motion is difficult to model because the forces driving the motion are difficult to predict. As a result these angles must be observed astronomically and made available to users operationally.

The last of the five angles characterizes the rotation angle of the Earth and is expressed as the time difference UT1–UTC (see Section 5.1.3). Principal variations in the rotation speed of the Earth include a constant deceleration due to tidal deceleration and deglaciation, decadal variations due to changes in the internal distribution of the Earth's mass, largely seasonal meteorologically driven variations, and tidally driven periodic variations. As with polar motion, UT1–UTC is difficult to model and predict, and must be observed astronomically and reported to users routinely. Figure 5.2 displays the motions and the reference system axes related to the changing orientation of the Earth in space.

5.1.1

Precession/Nutation

Gravitational forces exerted by solar system bodies act on the nonspherical Earth to cause its orientation in the celestial reference system to change. In 26 000 years the polar axis appears to describe a cone in space. This motion, called precession, is shown in Figure 5.3. Hipparchus is generally credited with discovering this effect in the second century BC, when he found that over time the positions of the equinoxes move westward along the ecliptic compared to the stars. The motion amounts to about 50 seconds of arc per year.



Figure 5.3 Schematic representation of precession and nutation.

Nutation refers to the much smaller periodic departures of the motion of the polar axis from the precessional motion. It, too, is due to gravitational forces acting on the Earth's bulge and was first detected by the English astronomer James Bradley (1748). The principal period of the nutational motion is 18.6 years, which is due to the lunar orbit.

In modern language, precession-nutation is the motion of the CIP in the GCRS, including the free core nutation and other corrections to the standard models. Precession is the secular part of this motion plus the term of 26 000-year period and nutation is that part of the CIP motion not classed as precession. To represent the motion of the CIP, mathematical models are used to specify the angular coordinates X and Y of the pole in the GCRS. In addition, a small correction obtained by astronomical observations can be applied. The IERS provides these 'celestial pole offsets' in the form of the differences, dX and dY , of the CIP coordinates in the GCRS with respect to the IAU 2000A precession-nutation model (i.e., the CIP is realized by the IAU 2000A precession-nutation plus these celestial pole offsets). In parallel the IERS also provides the offsets, in longitude and obliquity, with respect to the older IAU 1976/1980 precession/nutation model.

5.1.2

Polar Motion

Polar motion refers to the motion of the CIP with respect to the surface of the Earth and is made up principally of an approximately linear drift plus two periodic terms. The first of those periodic terms is a free motion of the pole, called the

Chandler wobble after its discoverer, Seth C. Chandler. Its period is 435 days and it corresponds to the free motion of a nonrigid Earth that was originally predicted by L. Euler (1765) for a rigid Earth. The second principal component of the polar motion is an annual motion driven by the seasonal redistribution of the Earth's atmospheric and water mass. Observational evidence for the existence of polar motion, which is seen as a variation of astronomical latitude, was gathered in the nineteenth century by astronomers and geodesists, including F. W. Bessel, F. G. W. Struve, C. A. F. Peters, O. W. Struve, J. A. H. Glydén, and M. Nyrén (Verdun and Beutler, 2000; Dick, 2000; Höpfner, 2000). Karl Friedrich Küstner is generally credited with establishing the reality of polar motion in the 1880s (Brosche, 2000; Küstner, 1888, 1890). Seth C. Chandler (1891a, 1891b, 1892) was the first to identify the major periodic components (Carter and Carter, 2000). Figure 5.4 shows the motion of the pole over the Earth's surface. This motion has been an active area of research since the nineteenth century (Munk and MacDonald, 1960; Lambeck, 1980; McCarthy, 2000).

Polar motion also includes sub-daily variations caused by ocean tides and periodic motions driven by gravitational torques with periods less than two days. It is specified by two angular coordinates of the pole with respect to the terrestrial system, x and y , as shown in Figure 5.4.

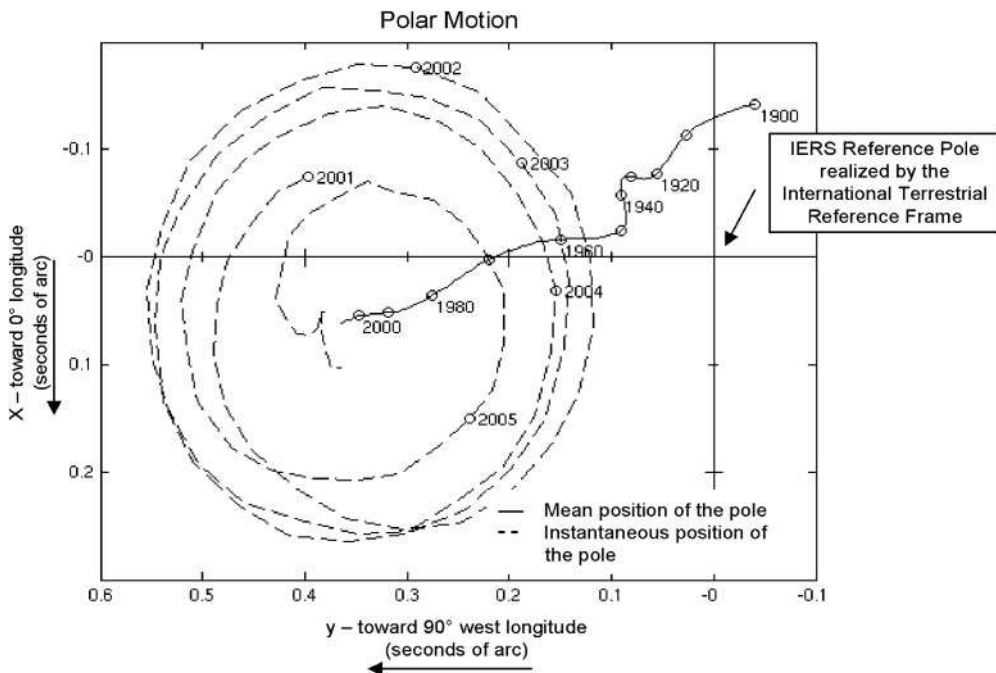


Figure 5.4 The trace of the pole over the Earth's surface from the International Earth Rotation and Reference System Services. The mean and instantaneous positions are shown separately.

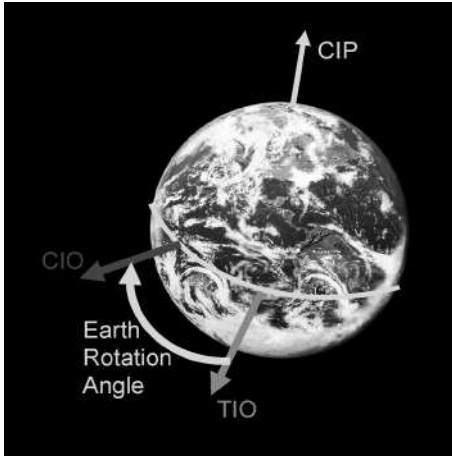


Figure 5.5 Earth rotation angle.

5.1.3 UT1

UT1 is a measure of the Earth's rotation angle expressed in time units and treated conventionally as an astronomical time scale defined by the rotation of the Earth with respect to the Sun. In practice, UT1 was defined until 1 January 2003 by means of a conventional formula (Aoki *et al.*, 1982). UT1 is now defined as being linearly proportional to the ERA, and the transformation between the ITRS and GCRS is specified using the ERA. The ERA is the angle measured along the intermediate equator of the CIP between the TIO and the CIO, positively in the retrograde direction and increasing linearly for an ideal, uniformly rotating Earth (Figure 5.5). It is related to UT1 by a conventionally adopted expression (5.1), in which ERA is a linear function of UT1.

$$\text{ERA} = \theta(T_u) = 2\pi(0.7790572732640 + 1.00273781191135448T_u), \quad (5.1)$$

where

$$T_u = (\text{Julian UT1 date} - 2451545.0)$$

$$\text{UT1} = \text{UTC} + (\text{UT1} - \text{UTC}).$$

Its time derivative is the Earth's angular velocity. UT1 is determined by astronomical observations (currently from VLBI observations of the diurnal motions of distant radio sources and other observations) and can be regarded as a time determined by the rotation of the Earth. It is obtained from the uniform time scale UTC by using the quantity UT1–UTC, which is provided by the IERS.

Greenwich Sidereal Time (GST) is an angle related by the equation of Aoki *et al.* (1982) to UT1. That expression has been updated to be consistent with recent IAU recommendations. GST is the sum of the ERA and the angular distance between the CIO and the equinox along the moving equator (this distance

corresponds to the accumulated precession and nutation in right ascension from the epoch of reference to the current date). GST contains both polynomial and periodic terms.

5.2

Variations in the Earth's Orientation

In a nonrotating system the torque equation is

$$\dot{\mathbf{H}} = \mathbf{L}, \quad (5.2)$$

where

\mathbf{L} is the torque

\mathbf{H} is the angular momentum.

In a reference frame rotating with angular velocity $\boldsymbol{\omega}$, this can be written as

$$\dot{\mathbf{H}} + \boldsymbol{\omega} \times \mathbf{H} = \mathbf{L}. \quad (5.3)$$

The angular momentum of a system of rotating particles can be expressed as the sum of two parts

$$\mathbf{H} = I\boldsymbol{\omega} + \mathbf{h}, \quad (5.4)$$

where I is the inertia tensor for matter in the volume v ,

$$I = \iiint_V r^2 \rho(\mathbf{r}) dV, \quad (5.5)$$

dV is the differential volume element with position $\mathbf{r} = (r_1, r_2, r_3)$ and spatial density $\rho(\mathbf{r})$. In (5.4), \mathbf{h} represents the relative angular momentum due to the motions $\mathbf{u}(\mathbf{r})$ relative to the system given by

$$\mathbf{h} = \iiint_V \rho(\mathbf{r})(\mathbf{r} \times \mathbf{u}) dV. \quad (5.6)$$

Equation (5.3) then becomes

$$\frac{d}{dt}(I\boldsymbol{\omega} + \mathbf{h}) + \boldsymbol{\omega} \times (I\boldsymbol{\omega} + \mathbf{h}) = \mathbf{L}, \quad (5.7)$$

which is referred to as the Liouville equation (Munk and MacDonald, 1960). It shows mathematically that torques, variations in the relative angular momentum, or any phenomenon that acts to modify the inertia tensor, will affect the angular velocity vector and consequently the orientation of the Earth. Precession-nutation and polar motion are described in the following sections, and the Earth's rotation angle variations are described in Section 4.5.

5.2.1

Precession Nutation

Knowledge of the orbits and masses of solar system objects, principally the Sun and Moon, allows us to determine the torques acting on the Earth. That infor-

mation along with an understanding of the internal constitution of the Earth permits us to derive mathematical expressions to model the precession-nutation motion of the Earth's axis in space. The IAU-recommended model for precession, called IAU 2006 Theory of Precession or P03 (Capitaine, Wallace and Chapront, 2003), provides the position of the CIP in the GCRS using six polynomial terms for both the X and Y components of the representation of this motion. In nutation, however, many more terms are required to represent the motion. The model recommended by the IAU, known as the IAU 2000A precession-nutation model (McCarthy and Petit, 2004) contains 1306 sin-cosine terms for X and 962 for Y, representing periodic motion components with periods ranging from hundreds of years to 2 days. It provides accuracy at the microarcsecond (μas) level. A shorter version using only 78 terms for each component provides accuracy at the single milliarcsecond (mas) level (McCarthy and Luzum, 2003).

Although these models provide a very accurate representation of the motion of the pole in the GCRS, astronomical observations indicate that small corrections, known as 'celestial pole offsets' are still observable. These are caused by the unmodeled motion of the Earth's core as well as other geophysical effects that are difficult to model.

5.2.2

Polar Motion

Polar motion refers to an extensive set of motions covering a range from secular to sub-daily timescales. Figure 5.6 shows the major components of the power spectrum, which are described in the following sections.

5.2.2.1 Secular Variation

Secular polar motion is the apparently aperiodic motion of the pole toward the direction of approximately 75° west longitude. For many years the reality of secular polar motion was in doubt because of questions regarding the possible systematic errors in the analysis of the astronomical observations. Recent observations, using modern, totally independent techniques, however, corroborate the past determinations, proving that this motion is real. The cause of secular polar motion is most likely the effects of post-glacial rebound. Other possibilities that have been raised include melting of glaciers and changes in continental water storage. The parameters of this component of polar motion are valuable sources of information, which are used to calibrate and check viscoelastic rebound models of the Earth's interior.

5.2.2.2 Decadal Variations

Decadal polar motion refers to the quasi-periodic motions of the pole with periods estimated to be between ten and forty years. The motions, with amplitudes of approximately 20 mas, are apparently real as they have been detected with both optical and independent modern techniques.

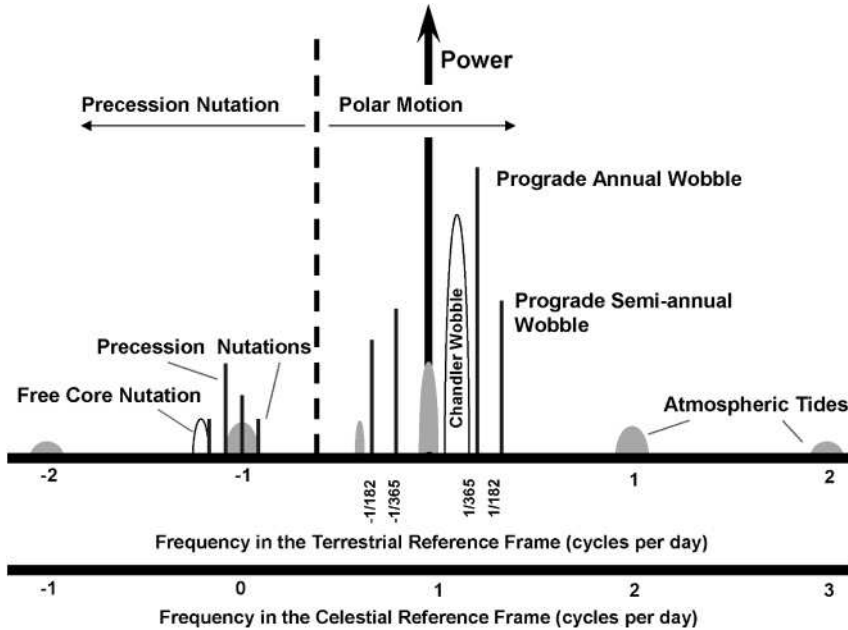


Figure 5.6 Major components of polar motion.

5.2.2.3 Chandler and Annual Variations

The power spectrum of polar motion (Figure 5.6) is dominated by the Chandler and annual components. Chandler motion refers to the motion of the CIP over the Earth's surface due to the fact that the rotation axis is not aligned with the axis of inertia. This motion was predicted by Euler (1765) for a rigid Earth, but first described observationally by Seth Chandler (1891a). The period for a rigid Earth would be approximately 306 days, but Chandler observed a period of 427 days. This discrepancy was explained by Simon Newcomb as being due to the fluidity of the oceans and the elasticity of the solid Earth (Newcomb, 1891). We now know that the Chandler motion is approximately circular with a period of approximately 433 days. We also know that the amplitude of the motion is variable, but that it is of the order of about 150 mas. The possible excitation of this motion remains mysterious. Possible causes include variations in pressure at the ocean bottom, variations in atmospheric pressure, groundwater, and seismic excitation, as well as core-mantle torque. Annual or seasonal polar motion is a stable prograde motion with an amplitude of approximately 90 mas. Causes are apparently atmospheric, oceanic, and groundwater excitations. The combination of the somewhat elliptical motion of the seasonal variation and the quasi-circular Chandler motion causes the spiral pattern seen in Figure 5.4.

5.2.2.4 Other Variations

The high-frequency portion of the polar motion spectrum contains a series of quasi-periodic motions with amplitudes of the order of one milliarcsecond. These

motions are poorly understood, but causes are assumed to be atmospheric and oceanic. Recent high-precision observations have also confirmed the existence of diurnal and sub-diurnal polar motion caused by ocean tides. They are reasonably modelable with amplitudes of the order of one milliarcsecond.

5.3

Transforming Between Reference Frames

Mathematically, the procedure to transform from a terrestrial reference system (TRS) to the celestial reference system (CRS) at the epoch t is written

$$[\text{CRS}(t)] = Q(t)R(t)W(t)[\text{TRS}(t)], \quad (5.8)$$

where $Q(t)$, $R(t)$ and $W(t)$ are the transformation matrices describing the motion of the precession-nutation, the rotation of the Earth around the axis of the pole, and polar motion, respectively (See McCarthy and Petit, 2004).

The parameter t , used in this and all of the following expressions, is defined by

$$t = [\text{TT} - 2000 \text{ January } 1, 12\text{h TT in days}] / 36\,525. \quad (5.9)$$

Note that 2000 January 1.5 TT = Julian Date 2451545.0 TT.

In the following discussion of the rotation matrices to be used in the transformations, we use the notation R_1 , R_2 , and R_3 to indicate rotations about the x , y , and z axes of the reference system respectively. That is:

$$\begin{aligned} R_1(\theta) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix}, \\ R_2(\theta) &= \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}, \\ R_3(\theta) &= \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \quad (5.10)$$

Referring to (5.8) the precession/nutation matrix $Q(t)$ can be written as

$$Q(t) = \begin{bmatrix} 1 - aX^2 & -aXY & X \\ -aXY & 1 - aY^2 & Y \\ -X & -Y & 1 - a(X^2 + Y^2) \end{bmatrix} \cdot R_3(s), \quad (5.11)$$

with

$$a = \frac{1}{2} + \frac{1}{8}(X^2 + Y^2). \quad (5.12)$$

X and Y are the angular ‘coordinates’ of the CIP in the CRS, and are provided by the conventional IAU 2000A or IAU 2000B precession-nutation models, based on

geophysical and astronomical theory. This formulation models the motion of the pole due to luni-solar and planetary motions. It does not include what has been called ‘planetary precession’ in the literature, which is the motion of the ecliptic caused by planetary gravitation. Before the IAU adopted this procedure in 2000 to describe the Earth’s precession and nutation, the ecliptic was used as a fundamental reference plane and its motion was described by a conventional expression based on astronomical theory. The combined precession-nutation of the Earth’s pole and the ecliptic was called general precession.

The quantity s specifies the position of the Celestial Intermediate Origin (CIO) on the equator of the Celestial Intermediate Pole (CIP). It is given mathematically by the expression

$$s = -XY/2 + \sum_i S_i t^j \sin(n_{il}l + n_{l'}l' + n_{F_i}F + n_{D_i}D + n_{\Omega_i}\Omega) + \sum_i C_i t^j \cos(n_{il}l + n_{l'}l' + n_{F_i}F + n_{D_i}D + n_{\Omega_i}\Omega) \quad (5.13)$$

where the coefficients S_i , C_i , n_i , n'_i , n_{F_i} , n_{D_i} , n_{Ω_i} are given in Table 5.1 for those terms with coefficients greater than $0.5 \mu\text{as}$, and

$$\begin{aligned} l &= \text{mean anomaly of the Moon} \\ &= 134.96340251^\circ + 1717915923.217800''t + 31.879200''t^2 \\ &\quad + 0.05163500''t^3 - 0.0002447000''t^4 \\ l' &= \text{mean anomaly of the Sun} \\ &= 357.52910918^\circ + 129596581.048100''t - 0.553200''t^2 + 0.00013600''t^3 \\ &\quad - 0.0000114900''t^4 \\ F &= L - \Omega \\ &= 93.27209062^\circ + 1739527262.847800''t - 12.751200''t^2 \\ &\quad - 0.00103700''t^3 + 0.0000041700''t^4 \\ D &= \text{mean elongation of the Moon from the Sun} \\ &= 297.85019547^\circ + 1602961601.209000''t - 6.370600''t^2 \\ &\quad + 0.00659300''t^3 - 0.0000316900''t^4 \\ \Omega &= \text{mean longitude of the ascending node of the lunar orbit} \\ &= 125.04455501^\circ - 6962890.543100''t + 7.472200''t^2 + 0.00770200''t^3 \\ &\quad - 0.0000593900''t^4 \end{aligned}$$

and where L is the mean longitude of the Moon.

The IAU has recommended that the precession-nutation model IAU 2000A, or the shorter version IAU 2000B for those who need a model accurate only to the level of one mas, be used to describe this motion. Software to implement these models can be found at http://tai.bipm.org/iers/conv2003/conv2003_c5.html.

Referring again to (5.8) the rotation of the Earth is given by

$$R(t) = R_3(-\theta), \quad (5.14)$$

θ being the Earth Rotation Angle at date t on the equator of the CIP.

Table 5.1 Polynomial coefficients used to determine s with coefficients greater than $0.5\mu\text{s}$. Units are given in micro-arcseconds.

i	j	S_i	C_i	n_I	n_I'	n_F	n_D	n_Ω
1	0	-2640.73	0.39	0	0	0	0	1
2	0	0.00	94.00	0	0	0	0	0
3	0	-63.53	0.02	0	0	0	0	2
4	0	-11.75	-0.01	0	0	2	-2	3
5	0	-11.21	-0.01	0	0	2	-2	1
6	0	4.57	0.00	0	0	2	-2	2
7	0	-2.02	0.00	0	0	2	0	3
8	0	-1.98	0.00	0	0	2	0	1
9	0	1.72	0.00	0	0	0	0	3
10	0	1.41	0.01	0	1	0	0	1
11	0	1.26	0.01	0	1	0	0	-1
12	0	0.63	0.00	1	0	0	0	-1
13	0	0.63	0.00	1	0	0	0	1
14	1	0.00	3808.65	0	0	0	0	0
15	1	-0.07	3.57	0	0	0	0	2
16	1	1.73	-0.03	0	0	0	0	1
17	2	743.52	-0.17	0	0	0	0	1
18	2	0.00	-122.68	0	0	0	0	0
19	2	56.91	0.06	0	0	2	-2	2
20	2	9.84	-0.01	0	0	2	0	2
21	2	-8.85	0.01	0	0	0	0	2
22	2	-6.38	-0.05	0	1	0	0	0
23	2	-3.07	0.00	1	0	0	0	0
24	2	2.23	0.00	0	1	2	-2	2
25	2	1.67	0.00	0	0	2	0	1
26	2	1.30	0.00	1	0	2	0	2
27	2	0.93	0.00	0	1	-2	2	-2
28	2	0.68	0.00	1	0	0	-2	0
29	2	-0.55	0.00	0	0	2	-2	1
30	2	0.53	0.00	1	0	-2	0	-2
31	3	0.00	-72574.11	0	0	0	0	0
32	3	0.30	-23.42	0	0	0	0	1
33	3	-0.03	-1.46	0	0	2	-2	2
34	4	0.00	27.98	0	0	0	0	0
35	5	0.00	15.62	0	0	0	0	0

It is obtained from the conventional relationship to UT1 (5.1),

$$\theta(T_u) = 2\pi(0.7790572732640 + 1.00273781191135448T_u), \quad (5.15)$$

where

$$T_u = (\text{Julian UT1 date} - 2451545.0)$$

$$UT1 = \text{UTC} + (\text{UT1} - \text{UTC})$$

or equivalently

$$\theta(T_u) = 2\pi \left(\frac{\text{UT1 Julian Days elapsed since } 2451545.0 + 0.7790572732640}{+ 1.00273781191135448T_u} \right). \quad (5.16)$$

Referring again to (5.8), the polar motion rotation is given by

$$W(t) = R_3(-s')R_1(\gamma)R_2(x). \quad (5.17)$$

The motion of the CIP within the TRF is not able to be modeled. Instead, this motion must be observed and accounted for appropriately in the transformation between coordinate systems. The size of these motions is small but very significant for precise transformation between reference frames. The polar coordinate information must be observed and reported. The data are available from the International Earth rotation and Reference System Service (IERS). They provide a series of files containing the latest data and predictions for the future. (see <http://www.iers.org/iers/products>.) s' can be approximated for the 21st Century as a function of time by

$$s' = -47 \mu\text{as T}. \quad (5.18)$$

5.4

Determination of Earth Orientation

Astronomical observations of polar motion have been carried out for over 100 years. A variety of instruments and techniques have been employed during that time. Optical observations of stars were first used to establish the existence of polar motion and to continue its monitoring. Transit circles, astrolabes, and zenith telescopes were the primary optical instruments, but both visual and photographic zenith telescopes were employed until the late 1980s as a significant source of polar motion information. Doppler observations of navigational satellites were used to improve determinations of polar motion during the 1970s and 1980s. Modern techniques came into existence in the 1970s when laser ranging to artificial Earth satellites, connected element and very long baseline interferometry (VLBI), began to make important contributions to our observational knowledge of polar motion. Currently, polar motion information is obtained principally from observations of the satellites of the Global Positioning System (GPS) and VLBI.

Observations of the Earth's rotation have been made for centuries using a host of astronomical instrumentation. Ancient civilizations used the direction of the Sun's shadow. The invention of the telescope was followed by its use to make observations of celestial bodies to measure time. As with polar motion, transit circles, astrolabes, and photographic zenith telescopes were used operationally for this purpose. Astronomical observations continue to be made to determine the difference between UT1 and UTC. VLBI currently provides the only true source of UT1–UTC information. Other space techniques, such as lunar laser ranging, laser ranging to satellites, and the determination of GPS satellite orbits, provide either an estimate of UT1 that depends on independent observations of polar motion using other techniques or the length of day, but not the actual Earth rotation angle.

Precession and nutation information historically was determined using transit circle observations. The most recent models, though, are based on VLBI observations of distant radio sources.

5.5 Earth Orientation Data

The International Latitude Service (ILS) was the first source of operational estimates of polar motion, using instrumentation consisting only of visual zenith telescopes. As more optical instrumentation became available the ILS was transformed into the International Polar Motion Service (IPMS), which included not only the visual zenith telescopes of the ILS, but also the visual and photographic telescope data of other international institutions. The Bureau International de l'Heure (BIH) began the routine collection of polar motion data to facilitate its mission to determine UT1. Determination of astronomical time (UT1) relied on accurate determination of polar motion, and the BIH collected and published polar motion information to provide a consistent set of UT1 and polar motion data. The optical data from 1899 to 1992 were analyzed for Earth rotation parameters by Vondrak (1999). The International Earth Rotation Service (IERS) began operation in 1988 following the recommendation of the MERIT Working Group. Essentially this service combined the operation of the IPMS and the BIH into one organization that made optimal use of the most accurate astronomical observations available to provide the ICRS, the ITRS, and the means to relate them to each other. Current polar motion and UT1 data are available from the International Earth Rotation and Reference System Service (IERS) at www.iers.org.

References

- Aoki, S., Guinot, B., Kaplan, G.H., Kinoshita, H., McCarthy, D.D. and Seidelmann, P.K. (1982) The new definition of universal time. *Astron. Astrophys.*, **105**, 359–61.
- Boucher, C., Altamimi, Z., Sillard, P. and Feissel-Vernier, M. (2004) *The ITRF2000*, International Earth Rotation Service Tech. Note 31, Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.
- Bradley, J. (1748) A letter to the Right Honourable George Earl of Macclesfield concerning an apparent motion observed in some of the fixed stars. *Philos. Trans.*, **45**, 1–43.
- Brosche, P. (2000) Küstner's observations of 1884–85: the turning point in the empirical establishment of polar motion, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 101.
- Carter, M. and Carter, W. (2000) Seth Carlo Chandler Jr.: the discovery of variation of latitude, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 109.
- Chandler, S.C. (1891a) On the variation of latitude, I. *Astron. J.*, **11**, 59–61.
- Chandler, S.C. (1891b) On the variation of latitude, II. *Astron. J.*, **11**, 65–70.
- Chandler, S.C. (1892) On the variation of latitude, VII. *Astron. J.*, **12**, 97–101.
- Capitaine, N., Wallace, P.T. and Chapront, J. (2003) Expressions for IAU 2000 precession quantities. *Astron. Astrophys.*, **412**, 567–86.
- Dick, S.J. (2000) Polar motion: a historical overview on the occasion of the centennial of the international latitude service, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 3.
- Euler, L. (1765) Du mouvement de rotation des corps solides autour d'un axe variable (On the movement of rotation of solid bodies around a variable axis). Originally published in *Mémoires de l'académie des sciences de Berlin*, **14**, 154–93.
- Höpfner, J. (2000) On the contributions of the Geodetic Institute Potsdam to the ILS, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 139.
- Kaplan, G.H. (2005) *The IAU Resolutions on Astronomical Reference Systems, Time Scales, and Earth Rotation Models; Explanation and Implementation*, U.S. Naval Observatory Circular 179, Washington, DC.
- Küstner, F. (1888) Neue Methode zur Bestimmung der Aberrations-Constante nebst Untersuchungen über die Veränderlichkeit der Polhöhe. *Beobachtungsergebnisse der Königlichen Sternwarte zu Berlin*, **3**, 1–59.
- Küstner, F. (1890) Über Polhöhen-Änderungen beobachtet 1884 bis 1885 zu Berlin und Pulkowa. *Astron. Nachr.*, **125**, 273.
- Lambeck, K. (1980) *The Earth's Variable Rotation*, Cambridge Press, Cambridge.
- Ma, C. and Feissel, M. (eds) (1997) *Definition and Realization of the International Celestial Reference System by VLBI Astrometry of Extragalactic Objects*, International Earth Rotation Service Tech. Note 23, Observatoire de Paris, Paris.
- McCarthy, D. (2000) Polar motion—an overview, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 223.
- McCarthy, D.D. and Luzum, B.J. (2003) An abridged model of the precession-nutation of the celestial pole. *Celest. Mech. Dynam. Astron.*, **85**, 37–49.
- McCarthy, D.D. and Petit, G.P. (eds) (2004) *IERS Conventions (2003)*, IERS Technical Note 32, Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.
- Munk, W.H. and Macdonald, G.J.F. (1960) *The Rotation of the Earth*, Cambridge University Press, London.

- Newcomb, S. (1891) On the periodic variation of latitude, and the observations with the Washington prime vertical transit. *Astron. J.*, **11**, 81–2.
- Perryman, M.A.C., Lindegren, L., Kovalevsky, J., Hog, E., Bastian, U., Bernacca, P.L., Creze, M., Donati, F., Grenon, M., Grewing, M., van Leeuwen, F., van der Marel, H., Mignard, F., Murray, C.A., Le Poole, R.S., Schrijver, H., Turon, C., Arenou, F., Froeschle, M. and Petersen, C.S. (1997) The hipparcos catalogue. *Astron. Astrophys.*, **323**, L49–52.
- Verdun, A. and Beutler, G. (2000) Early observational evidence of polar motion, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, CA, p. 67.
- Vondrak, J. (1999) Earth rotation parameters 1899.7:1992.0 after reanalysis within the hipparcos frame. *Surv. Geophys.*, **20**, 169–95.

6

Ephemeris Time

6.1

Need for a Uniform Time Scale

From the time of Newton the independent variable of time for all astronomical ephemerides and planetary theories had been assumed to be uniform. When it was established that the Earth's rotation and the mean solar time derived from the Earth's rotation were not uniform, the need arose for a new time scale to serve as the time argument for ephemerides and almanacs. Astronomers needed to provide a concept for an ideal time scale, a definition that approaches the ideal, and a practical realization that makes the time scale available.

In the 1930s and 1940s the clocks providing standard time were pendulum clocks. Quartz crystal clocks were coming into use, but there were concerns about their accuracy and reliability as the source of uniform time. There was a desire to introduce an astronomical time scale that was thought to be accurate, reliable, and uniform. Instead of using the rotation of the Earth, astronomers turned for a uniform time scale to the concept of a time-like argument based on the orbital motion of solar system bodies, and particularly to the apparent motion of the Sun that just reflects the orbital motion of the Earth. This new time scale, which came to be called Ephemeris Time, came into existence through a series of events. The definition of the new time scale would be given in terms of Newcomb's Theory of the Sun. The practical realization would be from observations of solar system bodies.

However, this step took over twenty years to come to completion. During that time there was significant progress in the development of new clocks based on developments in physics. The introduction and application of the theory of relativity for practical astronomical applications was extremely slow, because generally the accuracies did not require the complexities of the theory of relativity. This would change rapidly with the advent of artificial satellites and high speed digital computers.

6.2

Danjon Proposal

In 1948 the Comité International des Poids et Mesures (CIPM) referred a proposal to establish a uniform fundamental time standard to the International Astronomical Union (IAU). This proposal was considered at the International Colloquium on the Fundamental Constants of Astronomy in Paris in 1950 (*Colloque international sur les constantes fondamentales de l'astronomie*, 1950). There, suggestions were made that led to the establishment of Ephemeris Time (ET). However, in 1929 André Danjon, an astronomer at the Observatory of Strasbourg, had already written a paper, basically proposing the creation of a dynamical time scale. This paper and its contents apparently were either unknown, or forgotten, by the attendees of the Paris conference, with one possible exception. Danjon, who in 1950 was director of l'Observatoire de Paris and chairman of the meeting, made no reference to his earlier paper. Why he did this is unknown. It could have been because he felt the meeting was to make recommendations, and not review history. It could have been from modesty, or a sense of *noblesse oblige*.

In 1967 at an international colloquium on the gravitational problems of n bodies, (*Colloque sur le gravitationnel problème des n corps*, 1968), M. Jacques Lévy of l'Observatoire de Paris read a paper on the French contributions to the development of celestial mechanics in the last three centuries. He included a statement that can be translated as: 'It is A. Danjon who seems to have been the creator of the present ephemeris time; about 1929 he wrote "We should not delay longer in adopting a new practical definition of time. In order to fix the time by the position of the planets in their orbits it is necessary to make the calculations in celestial mechanics according to the law of Newton. A certain future date will be introduced in the calculations; the positions of the planets and the Moon will be obtained for this date; when observations show that these bodies pass through the calculated positions, the instant in question will be attained. It is thus that the bodies of the solar system mark the time in their orbits, graduated like a sun-dial".'

Further on in the paper Danjon continues, as translated by Clemence (1971):

Let us restrict ourselves to wishing that some day we may discover a good terrestrial standard of time, and leave these difficulties of pure logic, since also we are reassured about their practical repercussions. The dreaded objection falls before results of observation; no contradiction is to be feared, since the Moon and the planets give the same time. As the time so defined is furnished by the law of Newton, it is called Newtonian time; until experiments prove the contrary this time will be called absolute. How do we read the time on the dial of the planetary orbits?

There everything is yet to be done, or nearly everything. Of course we shall continue to employ terrestrial time derived from meridian passages of stars, perhaps as a first approximation, or perhaps as a means of interpolation, or rigorously for the needs of geodesy and navigation. But for the other applications,

celestial mechanics, the study of spectroscopic binaries and of short period variables, etc., we shall correct it by the differences between terrestrial time and Newtonian time, the latter being derived from the observations of the Moon, the Sun, and the interior planets.

Clemence did not discover this information until 1970 and wrote a paper on 'The concept of Ephemeris Time: A case of Inadvertent Plagiarism' (Clemence, 1971), from which the above information is taken.

6.3 Clemence's Proposal

In 1948 Clemence, in a paper on astronomical constants which included some pages on time measurement, wrote (Clemence, 1948):

The time used in the national ephemerides is the kind of time that satisfies the equations of celestial mechanics, often called Newtonian time, while the time used by astronomers in practice depends on the variable rotation of the earth. The use of one kind of time for the theory and another for the observations produces discrepancies between the observation and theory which are proportional to the mean motions of the objects observed ... the solar and planetary theories give the coordinates of the sun and planets as functions of Newtonian time, whereas the observed positions are functions of variable time; hence the theoretical coordinates differ from the observed. ... It therefore seems logical to continue the use of mean solar time ... for civil purposes, and to introduce Newtonian time for the convenience of astronomers and other scientists only Astronomers will continue to set their clocks to the mean solar time of the Greenwich meridian.... and to publish their observations on this standard of reference. Whenever an observation is to be compared with a theoretical position, the time of the observation will first be corrected to Newtonian time. The corrections to Universal time needed to reduce it to Newtonian time can be published in the national ephemerides.

In that paper, Clemence notes the earlier work of Spencer Jones (1939) in which he determined corrections to Newcomb's tabular longitude of the Sun (1895) that are required to match actual observations. These corrections are due to the fact that the Earth's variable rotational speed produces a nonuniform time scale that is not suitable for the precise dynamical description of the Sun's motion. He then went on to note that using these corrections has the effect of making the day defined by his 'Newtonian' time equivalent to a mean solar day of 1900. This is not correct, since Newcomb's Tables, while providing the solar longitude in terms for the year 1900, are actually based on much earlier observations of the Sun, and so the time scale in Newcomb's

work more closely resembles the mean solar second of the mid-nineteenth century (Nelson *et al.*, 2001).

At the International Colloquium on the Fundamental Constants of Astronomy in Paris in 1950, Clemence presented his paper and pointed out that an inconsistent mean solar second presented a problem to providing a consistent set of astronomical constants. He proposed that the national ephemerides routinely provide tabular values to apply to the astronomical Universal Time as observed (based on the mean solar time) in order to obtain what he called 'Universal Time (1900)' which he considered to be compatible with the mean solar second of 1900. He, therefore, proposed for discussion the proposition that the mean solar second of 1900 be used in all cases where the mean solar second of date was not appropriate. In the formal write-up of his paper in *Colloque international sur les constantes fondamentales de l'astronomie* (1950), Clemence acknowledges Spencer Jones for pointing out that what Clemence had called Universal Time (1900) is not the time at any one epoch but is more like the mean time during the nineteenth century. Spencer Jones in his paper offered the opinion that it would be better to 'specify the correction required to astronomical time, which is what is needed in investigations in dynamical astronomy, rather than to base everything on the length of the mean solar second at one particular epoch, because the data are not sufficiently accurate to permit of a precise specification of the length at a particular epoch.'

In discussions at that meeting, Clemence (1971) recalled that Spencer Jones made what Clemence considered to be the first suggestion that an atomic frequency along with a suitable means of integrating the measures might be used to provide time. Consistent with his 1929 paper, Danjon suggested that the motion of the Earth around the Sun might provide a more suitable definition of time than the solar day and proposed the term 'dynamical time' (Danjon, 1929). Brouwer proposed the name 'ephemeris time.'

Recommendation 6 of the conference read:

The Conference recommends that, in all cases where the mean solar second is unsatisfactory as a unit of time by reason of its variability, the unit adopted should be the sidereal year at 1900.0; that the time reckoned in these unit (*sic*) be designated *Ephemeris Time*; that the change of mean solar time to ephemeris time be accomplished by the following correction

$$\Delta t = 24.349s + 72.3165s T + 29.949s T^2 + 1.821 B,$$

where T is reckoned in Julian centuries from 1900.0 January 0 Greenwich Mean Noon and B has the meaning given by Spencer Jones in *Monthly Notices R.A.S.* (Vol. 99, 1939, p. 541) and that the above formula should also define the second (Spencer Jones, 1939).

No change is contemplated or recommended in the measure of Universal Time, nor in its definition.

6.4

Adoption and Definition

This 1950 Paris recommendation was adopted by the IAU Commission on Time (*Trans. Int. Astron. Union*, 1954) at the IAU General Assembly in Rome in 1952 with a slight modification of the numerical expression to read:

$$\Delta T = 24.349s + 72.318s T + 29.950s T^2 + 1.82144 B. \quad (6.1)$$

However, Danjon later pointed out that the tropical year is a better choice than the sidereal year, because, unlike the sidereal year, the tropical year's length is independent of the adopted value of precession. Therefore, the tenth meeting of the Conférence Générale des Poids et Mesures (CGPM) in 1954, following an earlier recommendation of the Comité International des Poids et Mesures (CIPM), proposed the following definition of the second: 'The second is the fraction 1/31 556 925.975 of the length of the tropical year for 1900.0.' (*Trans. Int. Astron. Union*, 1957). The numerical value of the fraction was based on Newcomb's formula for the geometric mean longitude of the Sun for the epoch of January 0, 1900, 12h UT (Newcomb, 1895) given by

$$L = 279^\circ 41' 46.04'' + 129\,602\,768.13'' T + 1.089'' T^2, \quad (6.2)$$

where T is the time reckoned in Julian centuries of 36 525 days.

From the value of the linear coefficient in Newcomb's formula, the tropical year of 1900 would then contain $[(360 \times 60 \times 60)/129\,602\,768.13] \times 36\,525 \times 86\,400 = 31\,556\,925.975$ s. This was subsequently adopted by the IAU at the General Assembly in 1955. Once again, Danjon offered a later comment, noting that the fraction ought to have a slightly more precise value in order to provide exact numerical agreement with Newcomb's formula. Consequently, in 1956 the CIPM adopted the slightly more precise value with the words: 'La seconde est la fraction 1/31 556 925.9747 de l'année tropique pour 1900 janvier 0 à 12 heures de temps des éphémérides.' The Comité also established the Comité Consultatif pour la Définition de la Seconde (CCDS) to coordinate the work of physicists on atomic standards and of astronomers on the astronomical standard of ephemeris time (*Procès Verbaux des Séances*, 1957). The ephemeris second, as a fraction of the tropical year, was formally adopted by the 11th CGPM in 1960. Finally, Ephemeris Time was adopted by the 10th General Assembly of the International Astronomical Union in Moscow in 1958 with the following statement: 'Ephemeris Time is reckoned from the instant, near the beginning of the calendar year A.D. 1900, when the geometric mean longitude of the Sun was $279^\circ 41' 48''.04$, at which instant the measure of ephemeris time was 1900 January 0d 12h precisely.'

With these definitions ephemeris time was equivalent to the system of time in Newcomb's Tables of the Sun. Thus, ephemerides determined from Newcomb's tables of the Sun and planets could be considered to have ephemeris time as the independent argument. The mean longitude of any planet, or the Moon, could have been used to define the epoch and rate of a uniform time system, and ephemeris time could be used as the independent variable in their tables and

theories. The definition of the ephemeris second was the first official definition of the second as a unit of measure.

6.5

Observational Determination

The recommendations provided a basis for a new, more uniform time scale, but its usefulness depended on the means to access this time scale. As we shall see, observations of stars using visual and photographic zenith tubes, transit circles, and astrolabes provided Universal Time. The standard civil time, given by broadcast signals and wire transmissions, was an intermediary measure of time that was adjusted as needed to agree with the observationally determined Universal Time, based on the variable rotation of the Earth. The difference between the two time scales, $\Delta T = ET - UT$, was determined by comparison of observations with ephemerides.

For dates before telescopic observations were available, observations of eclipses have been able to provide the long-term behavior of ΔT . For observations of total or near-total solar eclipses without accurate timing information it is still possible to determine ΔT . We can calculate the path of totality on the Earth's surface for any total solar eclipse in the past, assuming a constant rotational speed for the Earth. If an observer at that time is able to report that the eclipse actually occurred in a different location, we can interpret that geographical difference as being due to the fact that our assumption of the Earth's rotation angle, based on a constant rotational speed, was in error. The difference in the longitude between the predicted and actual locations of the eclipse observation can then be used to determine the value of ΔT that would be required to see the eclipse where it actually occurred. Figure 6.1 shows the method graphically.

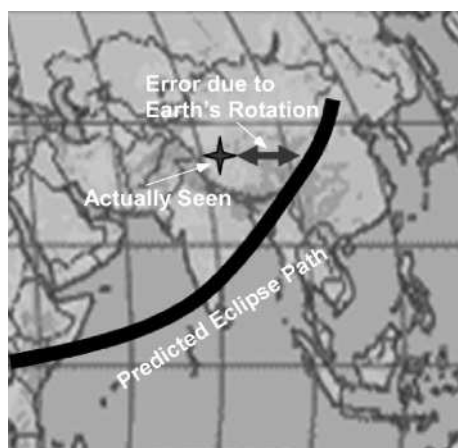


Figure 6.1 Determining ΔT from ancient eclipses.

In principle it would be possible to determine the position of the Sun in the sky with respect to a star catalog, determine the time in the adopted solar ephemeris when it should have been in that position, and then call that time the Ephemeris Time of the observation. This time could then be compared with the time determined from astronomical observations of the Earth's rotation angle to determine the corrections to be applied to the Earth's rotational time in order to provide the more uniform time scale. Although Ephemeris Time was based on the ephemeris of the Sun, observing the Sun directly is difficult, particularly with the precision needed to define a time scale. The faster moving, and more accurately observed, Moon was much more useful in determining ephemeris time from real observations.

Observations of the Moon with respect to star catalogs could be made using transit circles in the traditional manner. Timed observations of occultations of stars by the Moon also provided valuable ΔT data. A new instrument was developed in the 1950s by William Markowitz that was designed to determine positions of observing sites on the Earth, but found useful application in the determination of Ephemeris Time (Markowitz, 1954, 1961). It was designed to photograph the moon against a stellar background using a rotating filter to keep the position of the Moon constant during the exposure. Figure 6.2 shows the camera along with Markowitz. The observations provided a position of the Moon in the reference frame of the star catalog, whose star positions were used in the analysis of the lunar position.

Many difficulties remained with determining ephemeris time from the Moon. There were problems with the ephemeris of the Moon, specifically the secular acceleration value, the completeness of the gravitational theory of its motion, the irregularity of the limb profile, and the implied position of the equinox in the lunar theory.

Systematic differences in the positions of the equinox as realized in the theory of the Sun, the lunar theory, and the star catalogs became a problem, leading to



Figure 6.2 Markowitz Moon Camera.

a number of investigations, corrections, and papers concerning the equinox. For example, it was recognized that the origin in right ascension of the FK4 star catalog (Fricke and Kopff, 1963) was not the same at all declinations and different from the equinox of the almanac ephemerides of the Sun, the Moon, or numerically integrated ephemerides. The terms, ‘catalog equinox’ and ‘dynamical equinox’ were introduced to distinguish between the two types of equinoxes. Corrections were introduced for the equinox in the lunar theory to determine ephemeris time. Equinox position and motion corrections were introduced between the FK4 and FK5 (Fricke, Schwan and Lederle, 1988) star catalogs.

Observations of the Moon also required small corrections for the irregular surface features called ‘limb corrections.’ These features introduced errors in observations of the position of the Moon with transit circles and the timings of occultations of stars. These irregularities are responsible for the diamond ring and Bailey’s beads effects during solar eclipses. In 1963 C. B. Watts published his lunar limb charts, so that occultation observations could be corrected for the irregular limb of the Moon (Watts, 1963). These charts could provide corrections as large as 6 s in the determinations of Ephemeris Time.

In addition, Brown’s lunar theory (Brown, 1897–1908), the conventional basis for the lunar ephemeris of the time, could not be used directly to determine ephemeris time, because the theory was not strictly gravitational or completely in accord with Newcomb’s tables. This was corrected by removing the great empirical term from the lunar mean longitude and correcting the tabular mean longitude by

$$\Delta L = 8.72'' - 26.74'' T - 11.22'' T^2, \quad (6.3)$$

where T is measured in centuries of 36 525 ephemeris days from 1900.0.

Additional corrections were required for some of the periodic terms in longitude, latitude, and parallax. Beginning in 1960 the lunar ephemeris was calculated directly from an amended version of Brown’s theory, not from the tables. The *Improved Lunar Ephemeris* (1954) provided the information for dates from 1952 through 1959 using the same procedure. From 1960 to 1984 various improvements were introduced into the lunar theory. Different versions were designated as $j = 0, 1$, or 2 , and the corresponding ephemeris times were designated ET0, ET1, or ET2.

6.6

The Ephemeris Second and Atomic Time

Until 1960 the second was defined as $1/86\,400$ of the mean solar day, ignoring the variability in the Earth’s rotation and assuming that the Earth’s rotation was uniform. In 1960 the ephemeris second was introduced as the replacement for the second defined in terms of mean solar time. The ephemeris second was defined as $1/31\,556\,925.9747$ of the tropical year for 1900 January 0^d 12^h ET, as determined from Newcomb’s Theory of the Sun. This definition was more precise,

but it was difficult to realize from observations and to implement using operational clocks.

By 1960 Markowitz and his colleagues (Markowitz *et al.*, 1958) had already determined the value of the atomic second in terms of the ephemeris second. So in 1967 the *Système International* (SI) second was introduced as the duration of 9 192 631.770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium 133 atom. This numerical value is such that the SI second is equal to the value of the ephemeris second, as determined from the lunar observations from 1956–1965. This corresponds approximately to the mean solar second of the mid-nineteenth century. The ephemeris second had only lasted as the definition of the second for seven years.

The availability of operational atomic time scales, beginning in July 1955 (Essen and Perry, 1957), meant that ephemeris time could be determined from atomic time scales much more accurately than from solar system observations. Ephemeris time could, thus, be made readily available. The atomic time second had been defined to be equal to the ephemeris time second through the work of Markowitz *et al.* (1958). The actual realization of atomic time scales, however, was accomplished independently at a number of laboratories around the world. In 1967, the CCDS standardized atomic time scales by establishing the origin of International Atomic Time (TAI) to be in approximate agreement with 0^h UT2 on 1 January 1958 (*Com. Cons. Déf. Seconde*, 1970). In doing this, the CCDS followed the recommendation of IAU Commissions 4 (Ephemerides) and 31 (Time) in 1967. The result was that, since the atomic time second differed from the variable universal time second, atomic time immediately began to deviate from universal time. The offset between TAI, and Ephemeris Time, and its dynamical time successors is now established as 32.184 seconds. Consequently we can effectively extend the definition of Ephemeris Time into the era of atomic time using the relationship

$$ET = TAI + 32.184s. \quad (6.4)$$

The assumption is generally made that a time scale based on physical phenomena on Earth does not differ nonlinearly from a time scale based on solar system dynamics. We shall see that atomic clocks will become a source of accurate and reliable time and frequency, which will be the basis for many time scales and the most accurate physical measurements.

6.7 Historical ΔT

Historical records of the time-varying ΔT provide valuable information about long-term variations in the Earth's rotation. The observational data can be divided into the pre-telescopic (before 1600) and the telescopic observations (after 1600). In more recent times atomic clocks become available from 1955, and modern observations using lunar laser ranging and VLBI observations begin in 1969.

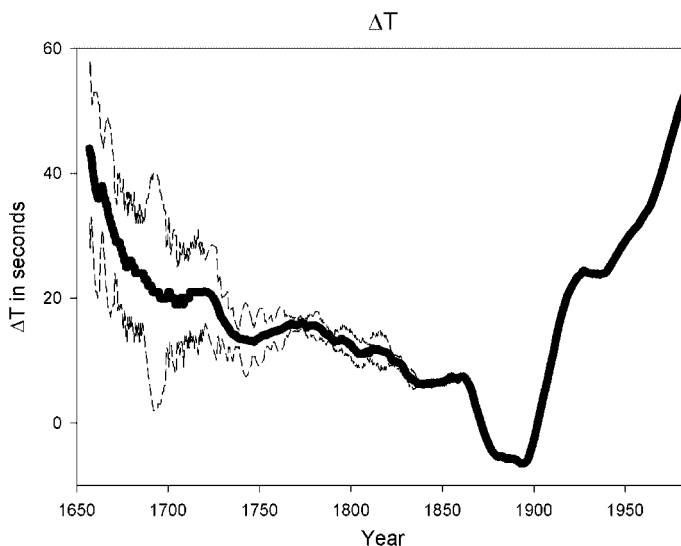


Figure 6.3 ΔT with error envelope from McCarthy and Babcock (1986).

For observations of the Moon the accuracy depends on the empirical value of the tidal secular acceleration in the mean longitude of the Moon included in the lunar ephemeris. Prior to lunar laser ranging data, Spencer Jones' value of $-22.4''/\text{cy}^2$ was used, modern investigations use a value of $-25.85''/\text{cy}^2$. Distinctions are also necessary concerning the lunar ephemerides used in the analyses. As we have seen, in the period 1960 to 1984 improvements were introduced into the lunar theory. Different versions of the lunar theory were designated by $j = 0, 1$, and 2 . The ephemeris times, determined from these theories, were designated as ET0, ET1, and ET2, respectively. The resulting values of ΔT could differ slightly depending on the ephemeris time used.

In 1952 Brouwer analyzed lunar occultation observations to determine values of ΔT from 1820 to 1950 (Brouwer, 1952). He detected a change in the length of the day of about 2 milliseconds per day per century during that interval and decadal fluctuations of several milliseconds per day. McCarthy and Babcock (1986) used lunar occultations for the early period and lunar laser ranging and VLBI observations when available to determine values of ΔT from 1657 to 1984. These data were smoothed to match the errors of the observations (Figure 6.3). Historical data back to 1000 BC have been compiled by L.V. Morrison and F. R. Stephenson, and these are shown in Figure 6.4.

Stephenson and Morrison have investigated extensively the historical values of ΔT back to ancient times. (Morrison and Stephenson, 2001, 2004, 2005; Stephenson, 1997, 2003; Stephenson and Morrison, 1984, 1994). In Stephenson and Morrison (1994) they arrived at conclusions that summarize their previous work and which may, in some instances, differ from their previous thoughts. They found that the average observed increase in the length of day over the previous

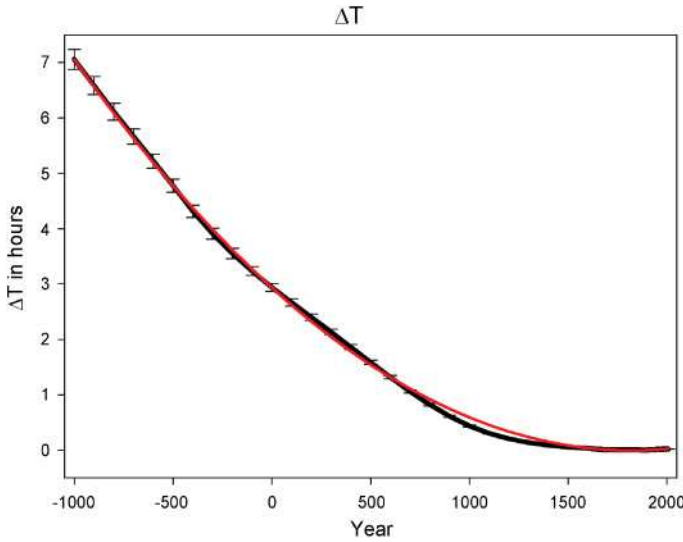


Figure 6.4 ΔT trend.

2700 years was $(1.70 \pm 0.05) \text{ mscy}^{-1}$, and that this was caused by two sources. The first, tidal dissipation, acts to increase the length of day by $(2.3 \pm 0.1) \text{ mscy}^{-1}$ which is commensurate with a tidal acceleration of the Moon of $-26.0'' \text{ cy}^{-2}$. A second, nontidal source acts to shorten the length of day by $(0.6 \pm 0.1) \text{ mscy}^{-1}$, a value consistent with modern values of the changing figure of the Earth and the theory of post-glacial rebound. They further found that there is an approximately 1500-year fluctuation with semi-amplitude 4 ms that may be due to interactions between the Earth's core and its mantle.

In Morrison and Stephenson (2004) the authors provide a table of values of ΔT at 10-year intervals from 1000 BC through 2000 along with their estimated errors. They further suggest that for low-precision applications the ΔT data can be represented by the parabola

$$\Delta T = -22 \text{ s} + 32 \text{ s} \times (\text{year} - 1820)^2. \quad (6.5)$$

That expression provides a gross means to extrapolate ΔT if only low levels of accuracy are required. The coefficient of the quadratic term in Eq. (6.5) corresponds to a change in the length of day of 1.75 milliseconds $\text{day}^{-1} \text{ century}^{-1}$, in close agreement with what we would expect to see from current geophysical theories. Figure 6.4 shows the data of Morrison and Stephenson (2004, 2005), McCarthy and Babcock (1986), and the most recent astronomical determinations. The colored line in the plot is the curve corresponding to Eq. (6.5). The departure of the observed ΔT data from the expression in Eq. (6.5) is shown in Figure 6.5 (Note the difference in units, hours and seconds, for the two figures.) It shows that significant departures from a strictly parabolic curve have occurred in the past,

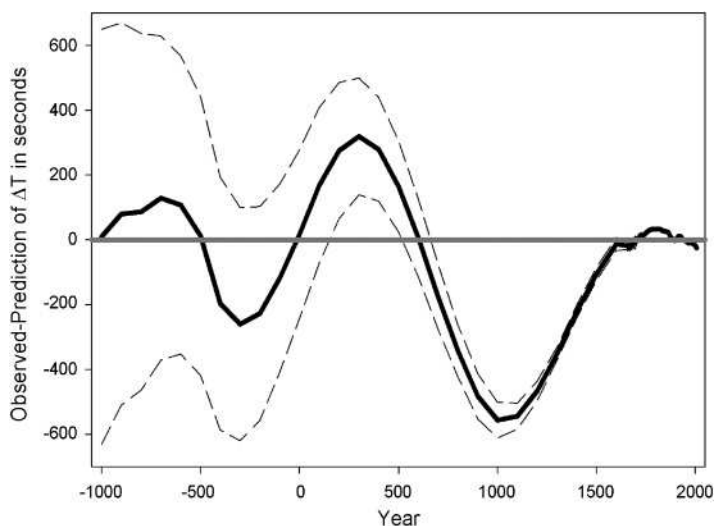


Figure 6.5 Departure from a parabolic fit to the ΔT trend with error envelope.

which indicates long-term variations in the Earth's rotation that do not fit a simple tidal deceleration model.

6.8

Problems with Ephemeris Time

Ephemeris Time was introduced as the time argument for ephemerides and astronomical almanacs in 1960. It was quickly apparent that atomic time provided a more accurate and available source of ephemeris time than solar system observations. It was also soon recognized that there were difficulties concerning the definition based on the theory of the Sun and the actual determinations based on observations of the Moon. The ephemerides also had problems accurately modeling the effects of tidal friction.

The definition of the second of ephemeris time was specified for 1900, based on Newcomb's theory, but the observational data to which the theory had been fitted dated from the 1700s and 1800s. So the ephemeris second corresponds roughly to the solar second for about 1820. In addition, the definition depends on values that were adopted by Newcomb for the aberration and precession constants. Hence, any improvements in the astronomical constants and reference systems would theoretically change the definition of ephemeris time. In practice, however, this was a conventionally adopted definition, and any such modifications to try to model a true mean solar second of 1900 were never considered.

There were no specified relationships between the versions of ephemeris time that could be determined from the observations of different solar system bodies, that is, the Sun, Moon, Mercury, or Venus. Each determination depended on the constants used in the ephemerides and the analyses of the observations. Inevitably,

further more accurate observations of the motions of solar system bodies would lead to improvements of the ephemerides and a resulting change in the realizations of ephemeris time.

Beside these technical problems with the definition and realization of Ephemeris Time there was a fundamental difficulty in the availability of the ephemeris second and ephemeris time for practical applications. There was, of necessity, a process of observation, analysis, and the comparison with ephemerides in order to determine either ephemeris time or the ephemeris second. This also meant that both could only be determined to a low level of precision. Atomic time would provide a much more available and precise time scale and second. At the same time that Ephemeris Time was introduced, atomic time scales were becoming more available, and the coordination of radio time signals using Universal Time was beginning to be more widespread. So the standard time scales were being steered, or stepped, to agree with Earth rotation times without consideration of the less available Ephemeris Time.

The timing of the introduction of Ephemeris Time in 1960 was additionally unfortunate, as this was the beginning of the space age, the introduction of digital computers, and a number of scientific developments requiring improvements in accuracies. Ephemerides based on numerical integrations, rather than analytical theories, were being developed, made possible by the computer capabilities becoming available. The astronomical constants and solar system theories of Newcomb were now over 60 years old and in need of improvement, which could, in theory, change ephemeris time.

Even more important, the theory of relativity had not yet really been applied to celestial mechanics and timekeeping, even though the accuracies being sought and achieved required the introduction of relativistic corrections. The equations of motion being used for numerical integrations of all the bodies of the solar system were beginning to include terms for relativity. It was recognized that the Lorenz transformations of general relativity distinguished between time scales at the solar system barycenter, the geocenter, and the Earth's geoid. Ephemeris Time was not specific as being barycentric or geocentric, and was used in both cases, without distinction. Ephemeris Time was neither a proper nor a coordinate time according to relativistic theory. At IAU colloquium 9 in Heidelberg, Germany, in August 1970, Irwin I. Shapiro, then of M.I.T., said 'The current definition of ephemeris time is philosophically repugnant, aesthetically horrifying, and completely inadequate' (Mulholland, 1972). However, it would take until 1976 for the International Astronomical Union to reach agreement to introduce improvements and changes. A number of changes became effective in 1984, including new dynamical time scales to replace Ephemeris Time.

6.9 Relativity

Prior to 1960 in much of the astronomical community the accuracies of ephemerides, star catalogs, time scales, and observations did not require the appli-

cation of the theory of relativity. So, although Einstein had published his paper on relativity in 1905, it had not been applied in most cases by the 1960s. In addition, there were some competing theories, and questions concerning the application of the theory. The result was the development of the parameterized post-Newtonian (PPN) formulations of the theory of relativity (Will, 1974). PPN introduced variable parameters that could be solved for (or set to) specific values according to the adopted assumptions. PPN was introduced by ephemerides developers in their numerical integrations in the 1960s and thereafter. Thus, they also had to introduce relativistic transformations for the reference frames and time scales of the observations.

The 1970s saw the formation of a number of IAU Working Groups to develop resolutions for changes in the reference system, fundamental star catalogs, planetary masses, astronomical constants, theory of nutation, and time scales. The decision was reached to introduce all the changes at the same time so there would be a clean break between the old and the new. The resolutions were adopted at the 1976 IAU General Assembly in Grenoble, France, and were to be introduced in astronomical publications in 1984 (Kaplan, 1982). Relativity was introduced, although with some reluctance to commit completely to the Einstein Theory of Relativity.

6.10 Dynamical Time Scales

The improving observational accuracy underlined the need to replace Ephemeris Time with other times that were conceptually based on solar system dynamics and the theory of relativity, properly defined, and realizable as the time arguments for ephemerides and almanacs. There was also the requirement to somehow extend time scales for times before 1955, when atomic time first became available, to be used in investigations of observations prior to 1955. Therefore, two new time-like arguments were introduced in 1984. These were Barycentric Dynamical Time (TDB) and Terrestrial Dynamical Time (TDT). The acronyms are given according to their French names. These time-like arguments were to be continuations of ephemeris time, but defined to correspond with their respective reference frames, either the geocenter or the barycenter. To achieve the continuity, TDT was specified with respect to TAI with an offset and rate to match ET. It was specified that TDB would only differ from TDT by periodic terms, so the epoch and rate of the two would agree. This specification would prove to be impossible to satisfy accurately, and name difficulties would arise, as seen in the next chapter.

References

- Brouwer, D. (1952) A study of the changes in the rate of rotation of the earth. *Astron. J.*, **57**, 125–46.
- Brown, E.W. (1897–1908) *Mem. R. Astron. Soc.* 1960, 48, Royal Astronomical Society, London.
- Clemence, G.M. (1948) On the system of astronomical constants. *Astron. J.*, **53**, 169–79.
- Clemence, G.M. (1971) The concept of ephemeris time: a case of inadvertent plagiarism. *J. Hist. Astron.*, **2**, 73–9.
- Colloque international sur les constantes fondamentales de l'astronomie. *Bull. Astron.*, **15**, 163–292, 1950.
- Com. Cons. Déf. Seconde, **5**, 21–23 (1970) Reprinted in *Time and Frequency: Theory and Fundamentals*, Natl. Bur. Stand. (U.S.) Monograph 140 (ed. B.E. Blair), U.S. Government Printing Office, Washington, DC.
- Comité international des poids et mesures (1957) Procès Verbaux des Séances, deuxième série, **25**, 77.
- Danjon, A. (1929) Le temps, sa définition pratique, sa mesure. *L'astronomie*, **43**, 13–22.
- Essen, L. and Perry, J.V.L. (1957) The cesium resonator as a standard of frequency and time. *Philos. Trans. R. Soc. London A*, **250**, 45–69.
- Fricke, W., Kopff, A., in collaboration with Gliese, W., Gondolatsch, F., Lederle, T., Nowacki, H., Strobel, W. and Stumpff, P. (1963) Fourth Fundamental Catalogue (FK4). *Veröff. Astron. Rechen-Inst. Heidelberg*, **10**, 1–144.
- Fricke, W., Schwan, H., Lederle, T. in collaboration with Bastian, U., Bien, R., Burkhardt, G., du Mont, B., Hering, R., Jährling, R., Jahreis, H., Röser, S., Schwerdtfeger, H.M. and Walter, H.G. (1988) Fifth Fundamental Catalogue (FK5). Part I. The basic fundamental stars. *Veröff. Astron. Rechen-Inst. Heidelberg*, **32**, 1–106.
- Improved Lunar Ephemeris 1952–1959*, U.S. Government Printing Office, 1954.
- Kaplan, G.H. (ed.) (1982) The IAU resolutions on astronomical constants, time scales, and the fundamental reference frame, in *U.S. Naval Observatory Circular* 163, USNO, Washington.
- Markowitz, W. (1954) Photographic determination of moon's position, and applications to the measure of time, rotation of the earth, and geodesy. *Astron. J.*, **59**, 69–73.
- Markowitz, W. (1961) The photographic zenith tube and the dual-rate moon-position camera, in *Telescopes. Stars And Stellar Systems* (eds G.P. Kuiper and B.M. Middlehurst), University of Chicago Press, Chicago, IL, p. 88.
- Markowitz, W., Hall, R.G., Essen, L. and Perry, J.V.L. (1958) Frequency of cesium in terms of ephemeris time. *Phys. Rev. Lett.*, **1**, 105.
- McCarthy, D.D. and Babcock, A. (1986) The length of day since 1656. *Phys. Earth Planet. Inter.*, **44**, 281–92.
- Morrison, L.V. and Stephenson, F.R. (2001) Historical eclipses and the variability of the earth's rotation. *J. Geodynamics*, **32**, 247–65.
- Morrison, L.V. and Stephenson, F.R. (2004) Historical values of the earth's clock error ΔT and the calculation of eclipses, *J. His. Astron.*, **35**, 327–36.
- Morrison, L.V. and Stephenson, F.R. (2005) ADDENDUM historical values of the earth's clock error. *J. Hist. Astron.*, **36**, 339.
- Mulholland, J.D. (1972) Measures of time in astronomy. *Pub. Astron. Soc. Pacific*, **94**, 357–64.
- Nelson, R.A., McCarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L. and Bartholomew, T.R. (2001) The leap second: its history and possible future. *Metrologia*, **38**, 509–29.
- Newcomb, S. (1895) *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac*, Vol. VI, Part I: *Tables of the Sun*, U.S. Government Printing Office, Washington, DC, p. 9.
- Spencer Jones, H. (1939) The rotation of the earth and the secular acceleration of the sun, moon, and planets. *Mont. Not. R. Astron. Soc.*, **99**, 541.
- Stephenson, F.R. (1997) *Historical Eclipses and Earth's Rotation*, Cambridge University Press, Cambridge.

- Stephenson, F.R. (2003) Harold Jeffreys Lecture 2002; historical eclipses and earth's rotation. *Astron. Geophys.*, **44**, 2.22–2.27.
- Stephenson, F.R. and Morrison, L.V. (1984) Long term changes in the rotation of the Earth: 700 B.C. to A.D. 1980, in *Rotation in the solar system* (ed. R. Hide), Royal Society, London; reprinted in *Philos. Trans. R. Soc. London A*, **313**, 47–70.
- Stephenson, F.R. and Morrison, L.V. (1994) Long-term changes in the rotation of the Earth: 700 B.C. to A.D. 1990. *Philos. Trans. R. Soc. London*, **351**, 165–202.
- Trans. Int. Astron. Union*, Vol. VIII, *Proceeding of 8th General Assembly, Rome, 1952* (ed. P.T. Oosterhoff), Cambridge University Press, New York, p. 66, 1954.
- Trans. Int. Astron. Union*, Vol. IX, *Proceeding of 9th General Assembly, Dublin, 1955* (ed. P.T. Oosterhoff), Cambridge University Press, New York, p. 451, 1957.
- Trans. Int. Astron. Union*, Vol. X B, *Proceeding of 10th General Assembly, Moscow, 1958* (ed. D.H. Sadler), Cambridge University Press, 1960.
- Watts, C.B. (1963) The marginal zone of the moon. *Astron Papers for the Astronomical Ephemeris and Nautical Almanac*, 17, U.S. Government Printing Office.
- Will, C.M. (1974) *Experimental Gravitation* (ed. B. Bertotti), Academic Press, New York.
- Zanichelli, N. (ed.) (1968) *Colloque sur le gravitationnel problème des n corps, Paris, France, 16–18 Août, 1967. (Conférence on the Gravitational n-Body Problem, Paris, France, August 16–18, 1967)*, Publication du Centre National de la Recherche Scientifique, Paris.

7

Relativity and Time

7.1

Newtonian Reference Systems

While Einstein's theory of relativity was published in 1905 and a number of consequences of the theory were predicted, fifty years later it was not being applied for timekeeping or reference systems. The accuracies at that time just did not require the complications of the theory. The Newtonian gravitational theory was adequate, and it was easily understood. Special relativity was applied in restricted circumstances. General relativity goes to the basics of time and coordinate systems.

However, in the 1960s the accuracies were improving because of technology developments and the requirements of the space age. In 1976 the International Astronomical Union introduced relativistic concepts for time and the transformations between time scales. Reference systems based on the theory of relativity were not introduced until 1991.

In Newtonian systems we accept the Newtonian universal law of gravity and mechanics. Time is an independent absolute progression in terms of some unit of time. The coordinate system is in the framework of Euclidean geometry. All systems of coordinates are equivalent or interchangeable, as long as they move with a constant velocity. Transformations between time scales and coordinate systems are linear functions or angular rotations. The concepts are generally easily visualized.

7.2

Special Relativity

In 1905 Einstein proposed the theory of special relativity to explain several effects that seemed to contradict Newtonian physics and mechanics. It deals with motions in inertial reference frames. A reference frame is inertial if a free particle that is not subject to any external force moves without acceleration with respect to the frame. In Newtonian mechanics, such free particles move with constant speed along straight lines. In the Galilean principle of relativity, if a reference frame is

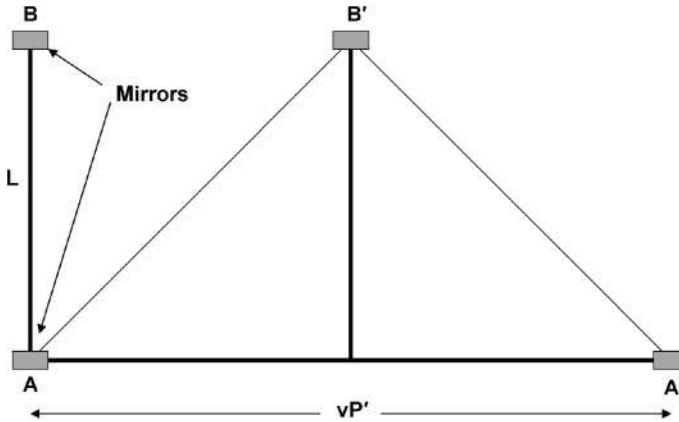


Figure 7.1 Time dilation.

inertial, all reference frames that move with a constant speed with respect to that reference frame are also inertial. Einstein generalized this principle by stating that all inertial frames are totally equivalent for the performance of all physical experiments.

In 1887 Michelson and Morley tested the law of addition of velocities. It was thought that light was propagated through ether, so a test was devised to determine the orbital velocity of the Earth with respect to the ether by measuring the velocity of light along, opposite to, and perpendicular to the Earth's motion. The speed was found to be the same in all directions. The conclusion was that there was no ether and no velocity could be larger than the speed of light. Thus, light is seen to be traveling with a constant speed, whether the source and/or the observer are fixed or moving. The independence of light velocity on the motion of the source can also be demonstrated using binary star observations.

This development has implications for precise timing. Referring to Figure 7.1, if we have mirrors at A and B separated by length L , in a reference system (S) and a light pulse repeatedly reflected between them, an observer at A with a clock would expect to measure the periodic return of the pulse with a period, P , equal to the distance the light pulse has to travel divided by the speed of light, or $P = 2L/c$, where c is the speed of light in a vacuum.

Another observer in a system (S') moving with constant speed v perpendicular to the line connecting the mirrors will also observe a periodic response with a period equal to the distance the light travels divided by the speed of light. However, in this case, we would expect the light to return after being reflected at B' to A' . Applying the Pythagorean Theorem, the distance traveled will be

$$d = 2\sqrt{\left(\frac{vP'}{2}\right)^2 + L^2}. \quad (7.1)$$

We would then expect to observe a response with period

$$P' = \frac{d}{c} = \frac{2}{c} \sqrt{\left(\frac{vP'}{2}\right)^2 + L^2}. \quad (7.2)$$

From this we find

$$P' = \frac{2L}{c} \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (7.3)$$

Then the apparent period for the moving observer is $P' = P\gamma$, with

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (7.4)$$

Thus, the moving clock ticks more slowly than the fixed clock by $1/\gamma$. Because of the fixed speed of light, a time interval measured in a moving frame is longer than the same time interval measured in a frame fixed with respect to the events. This phenomenon is known as time dilation. A similar fundamental result is the contraction of moving objects. A rod in motion relative to an observer will have its length contracted along the direction of motion as compared to an identical rod that is stationary. If L is the length of the rod at rest, the apparent length of the rod moving with velocity v is $L' = L/\gamma$.

7.3

Lorentz Transformations

The relationships above were formalized by Lorentz into transformations between spatial coordinates \mathbf{r} (x, y, z) and time t in a reference frame S and \mathbf{r}' (x', y', z') and time t' in a reference frame S' moving with a constant velocity \mathbf{V} with respect to S .

The vector form of the Lorentz transformation is

$$\begin{aligned} t' &= \gamma(t - \mathbf{V} \cdot \mathbf{r}/c^2) \\ \mathbf{r}' &= \mathbf{r} - \left[\gamma t - (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{r}}{|\mathbf{V}|^2} \right] \mathbf{V}, \end{aligned} \quad (7.5)$$

where γ is given by (7.4). The inverse formulae are

$$\begin{aligned} t &= \gamma(t' + \mathbf{V} \cdot \mathbf{r}'/c^2) \\ \mathbf{r} &= \mathbf{r}' + \left[\gamma t' + (\gamma - 1) \frac{\mathbf{V} \cdot \mathbf{r}'}{|\mathbf{V}|^2} \right] \mathbf{V}. \end{aligned} \quad (7.6)$$

If we assume the velocity V is along the x axis, $V = (v_x, 0, 0)$, then the y and z components are identities. The terms independent of γ vanish and (7.5) becomes

$$\begin{aligned} t' &= \gamma \left(t - \frac{v_x x}{c^2} \right) \\ x' &= \gamma (x - v_x t) \\ y' &= y \\ z' &= z \end{aligned} \tag{7.7}$$

the form in which the Lorentz transformation is usually given.

The space coordinates are x , y , z , and ct has the dimension of length. t is then called coordinate time. So in special relativity, and in general relativity, an event is specified in four coordinates, three spatial and one time coordinate.

If Δt and Δx , Δy , Δz are the coordinate differences between two events, P and Q, the quantity Δs^2 defined by

$$\Delta s^2 = -c^2 \Delta t^2 + \Delta x^2 + \Delta y^2 + \Delta z^2 \tag{7.8}$$

is invariant under the change of coordinates. Δs is the interval between the two events, and (7.8) defines the Minkowski metric. This shows that the interval is a quantity that can be associated with two events without referring to any particular coordinate system. The quantity Δs^2 is a generalized distance that is called the invariant interval. It is usually given as an infinitesimal displacement and written as

$$ds^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2. \tag{7.9}$$

7.4

Coordinate and Proper Time

In special and general relativity there are two types of time, proper and coordinate. Proper time is measured along the trajectory of an observer in space time ('world-line'). In practice, it is measured by a physical clock accompanying the observer. The clock must be insensitive to environmental conditions, gravity, and accelerations. Proper time is invariant in any coordinate change. Since the SI second is defined only in terms of the periods of radiation for the cesium atom, it contains no indication of a specific gravitational potential, or state of motion. Thus, any observer can realize the SI second as the unit for proper times. Proper time cannot be used to describe phenomena in extended domains, in which cases coordinate time must be used.

In special and general relativity, a four-dimensional space time reference system uses three spatial coordinates (x_1 , x_2 , x_3) and a fourth, $x_0 = ct$, where t is the coordinate time in this reference system. Coordinate time is an unambiguous way of dating in a specific reference system and is to be used as the time basis in the theory of motion in the system. In metrology it can be argued that coordinate time

cannot be measured, but only computed. The relation of proper time of an observer to coordinate time is provided by the metric, which takes into account the surrounding masses and energy.

If two events occur separated by a time dt at the same place in a reference frame, the interval (7.9) reduces to

$$ds^2 = -c^2 dt^2. \quad (7.10)$$

The time t is linked with the place and the new time is called the proper time, τ , which describes the local physics of the point. Then the interval (7.9) can be written for the general case as

$$ds^2 = -c^2 d\tau^2 = -c^2 dt^2 + dx^2 + dy^2 + dz^2. \quad (7.11)$$

This means that for the interval s^2 between two events, the quantity $\sqrt{s^2}/c$ equals the difference in readings of a clock moving at a constant velocity between two events. To derive the relationship between proper and coordinate time, consider the components of the constant velocity as measured with respect to coordinate time.

$$c^2 \left(\frac{d\tau}{dt} \right)^2 = c^2 - \left[\left(\frac{dx}{dt} \right)^2 + \left(\frac{dy}{dt} \right)^2 + \left(\frac{dz}{dt} \right)^2 \right], \quad (7.12)$$

from which

$$\frac{d\tau}{dt} = \sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{\gamma}. \quad (7.13)$$

This is the same expression as that shown in Section 7.2 for the time dilation relationship between clocks in reference frames moving with a velocity v with respect to each other. So for the moving observer the period of the clock is in coordinate time, while the fixed observer measures in proper time. For an observer at rest in the reference frame, the proper time coincides with the coordinate time in that reference frame in special relativity. This is not the case for observations of moving objects. This concept is visualized with the Minkowski diagram.

7.5 Minkowski Diagrams

The Minkowski diagram was developed in 1908 by Herman Minkowski and it provides an illustration of the properties of space and time in the special theory of relativity (Minkowski, 1908). It allows a quantitative understanding of the corresponding phenomena like time dilation and length contraction without mathematical equations. It is a space-time diagram with usually only one spatial dimension. Referring to Figure 7.2, distance is displayed on the x -axis and time on the y -axis. Events happening on a horizontal path in space can then be represented on a horizontal line in the diagram. Each point in the diagram represents an event in space and time, and the curve it follows is called its

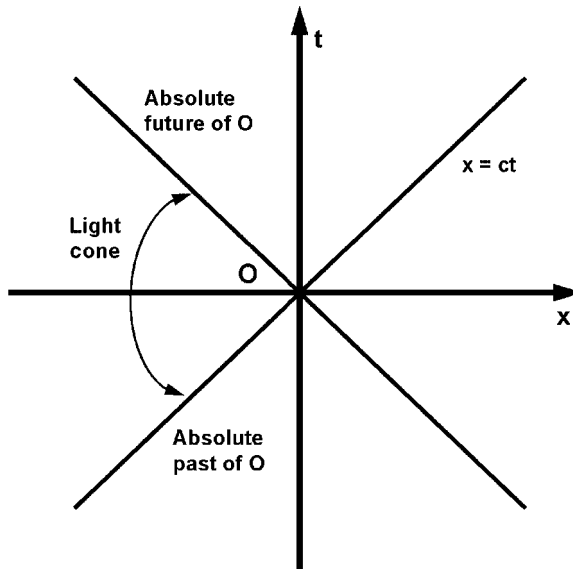


Figure 7.2 Minkowski diagram.

world line. It is called an event whether or not anything actually happens at that position. Objects plotted on the diagram can be thought of as moving from bottom to top as time passes, so a vertical line corresponds to a fixed point in space.

The world lines of light pulses are represented by straight lines defined by $x = ct$. The world line of a photon traveling through the origin to the right is a straight line with a slope of 45° if the scales on both axes are chosen appropriately. If light cones are drawn in the positive and negative time directions from a certain event (O), space time is separated into distinct 'future' and 'past' regions. The future is the locus of all events that have not yet happened and can be affected. An event cannot affect anything outside of its future light cone, because, in order to do so, it would have to send some sort of message to the desired location faster than the speed of light. The past is the locus of events that contributed to the current state. Anything that happened before and is not in the past light cone could not have affected the present, because its future light cone does not encompass the present.

We can superpose the coordinate systems for two observers moving relative to each other with constant velocity v in a Minkowski diagram to allow, the space and time coordinates, x and t , used by one observer, to be read off immediately with respect to the corresponding x' and t' used by the other, and vice versa (Figure 7.3). Both observers would assign an event at A to different times and locations. Events which are estimated to happen simultaneously from the viewpoint of one observer happen at different times for the other. In the Minkowski diagram this simultaneity corresponds with the introduction of a separate path axis for the moving observer. Each observer interprets all events on a line parallel

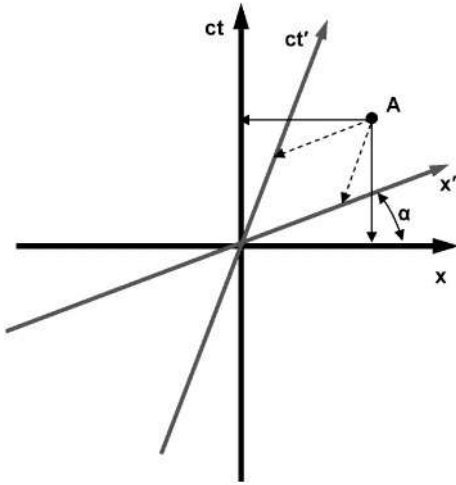


Figure 7.3 Minkowski diagram showing the superposition of two coordinate systems moving relative to each other with a constant velocity.

to his path axis as simultaneous. The sequence of events from the viewpoint of an observer can be illustrated graphically by shifting this line in the diagram from bottom to top. If ct is assigned on the time axes, the angle α between both path axes will be identical with that between both time axes. This follows from the speed of light being the same for all observers, regardless of their relative motion. α is given by

$$\tan \alpha = \frac{v}{c}, \quad (7.14)$$

where v is the relative velocity between the reference frames. The corresponding transformation from x and t to x' and t' and vice versa is described mathematically by the Lorentz transformation.

Time dilation is illustrated in Figure 7.4. In this case both observers consider the clock of the other as running slower. Relativistic time dilation means that a clock moving relative to an observer is running slower, and also the time in this system. This can be read immediately from Figure 7.4. The observer and clock at A is assumed to move from the origin O toward A, and the clock from O to B. For the observer at A all events happening simultaneously in this moment are located on a straight line parallel to its path axis passing A and B. Since $OB < OA$ the observer concludes that the time passed on the clock moving relative to him is less than that passed on his own clock, since they were together at O. A second observer having moved together with the clock from O to B will argue that the other clock has reached only C, and, therefore, this clock runs more slowly. The reason for the apparently paradoxical statements is the different determination of the events

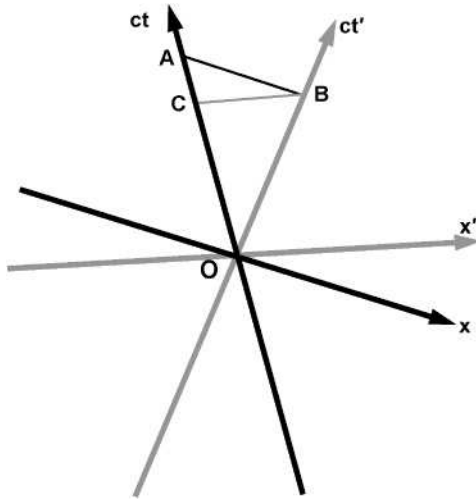


Figure 7.4 Minkowski diagram illustrating time dilation.

happening synchronously at different locations. According to the principle of relativity, the question of who is right has no answer and does not make sense.

7.6

Time in Special Relativity

The coordinate time in a reference frame must be given by a clock that is fixed relative to the reference frame. The second viewed as the unit of proper time is independent of the frame. So the synchronization of clocks in the frame must involve clocks fixed with respect to that frame. Clocks fixed with respect to the frame are synchronized according to Einstein's synchronization rule with respect to clocks in a moving frame because of time dilation. Hence, corrections are necessary to compare clocks with different velocities.

Likewise an observer in a fixed frame observing a frequency from a source in a frame moving with velocity \mathbf{v} will detect a frequency difference due to the Doppler Effect. The difference between the emitted frequency, f , measured in the source's frame and the received frequency, f' , measured in the observer's frame is

$$f' = f \frac{\left(1 + \frac{\mathbf{v} \cdot \mathbf{n}}{c}\right)}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (7.15)$$

where \mathbf{n} is the unit vector along the straight path of the signal in the direction of the propagation.

7.7

General Relativity

While special relativity applies to an empty and, hence, unreal universe, general relativity is a model of the real universe, where mass and energy find their proper place. General relativity is a theory of gravity. Unlike the fixed framework of Newtonian space time, here gravitational effects are not viewed as action at a distance; instead they appear as local geometrical properties of a Riemannian space time, curved by the presence of energy and mass.

In Einstein's Theory of Relativity, the inertial frame is generalized by the concept of the free-falling isolated local frame, which is electrically and magnetically shielded, sufficiently small that inhomogeneities in the external fields can be ignored throughout the volume, and in which self-gravitating effects are negligible. In this free-falling frame any local nongravitational test experiment is independent of where and when in the universe it is performed and of the velocity of the frame.

In general relativity there are no privileged coordinate systems, although there are some that are more convenient than others. The geometrization of the gravitational field is a formulation of a general principle of general relativity, the equivalence principle.

7.7.1

Metrics in General Relativity

In Newtonian mechanics, an object located at a point \mathbf{P} with coordinates (x, y, z) in a three-dimensional Cartesian reference frame, in the presence of a gravitational potential $U(x, y, z)$, would experience a force with components of

$$\frac{\partial U}{\partial x}, \frac{\partial U}{\partial y}, \frac{\partial U}{\partial z}. \quad (7.16)$$

So the force, \mathbf{f} , at point, \mathbf{P} , can be related to U by

$$\mathbf{f} = \nabla U. \quad (7.17)$$

If there are N masses, m_i , located at \mathbf{Q}_i acting on the point, \mathbf{P} , the Newtonian potential is

$$U = G \sum_{i=1}^N \frac{m_i}{|\mathbf{P} - \mathbf{Q}_i|}, \quad (7.18)$$

where G is the universal constant of gravity. The potentials for finite gravitating bodies, the spherical harmonic development for point outside the sphere, the additional terms for relativity, and gravito-magnetic fields can be found in Kovalevsky and Seidelmann (2004).

For general relativity, the general four-dimensional metric is

$$ds^2 = \sum_{i=0}^3 \sum_{j=0}^3 g_{ij} dx_i dx_j, \quad (7.19)$$

where $g_{ij} = g_{ji}$ and the index 0 corresponds to the time coordinate. In general relativity, if the potential vanishes and there is no spatial rotation and no acceleration of the origin, the expression reduces to the space time special relativity metric.

$$ds^2 = -c^2 dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2. \quad (7.20)$$

The coefficients g_{ij} can be expanded as a function of small parameters. To the accuracies of microarcseconds in direction and 10^{-17} s in time for a weak field like the solar system, the IAU in 2000 recommended expansions as given in the resolutions in Section 7.8.

7.7.2

The Equivalence Principle

In classical mechanics there are two distinct types of matter. Inertial mass m' is the mass appearing in Newton's second law: $F = m'a$; the inertial mass is the measure of inertia of a body, a measure of how difficult it is to change its velocity. The gravitational mass m is the mass appearing in the expression for Newtonian attraction: $F = G m_1 m_2 / r^2$. Gravitational mass is the measure of the magnitude of the gravitational force F generated by a body. To the accuracy of observations, it has long been recognized that $m = m'$, and this weak equivalence principle is used in classical mechanics.

Einstein generalized this weak equivalence principle to a free-falling laboratory near a gravitating mass being equivalent to a laboratory in free motion, as on an artificial satellite or a space probe outside the solar system. Generally, the equivalence principle states that there is no experiment that allows one to distinguish whether a laboratory is in a uniformly accelerating elevator, or if it is a fixed laboratory in a gravitational field with a uniform strength over the space of the laboratory.

Thus, there is no preferred local inertial reference frame. All are equivalent and all nongravitational laws of nature take the same form as in special relativity. Thus, for a locally inertial reference frame, there is a metric that has the form and properties of the special relativity metric describing a flat four-dimensional space time. In general relativity there is an extended metric, which depends on the distribution and motion of masses in the universe, and at each point there is a flat special relativity space time, tangent to the general relativity space time such that the four coordinates and their derivatives with respect to the two space times can be set to correspond. In general relativity space time, a test particle moves on a geodesic of the metric. So light bends in a gravitational field because the geodesic is not a straight line. Also, light traveling through a gravitational field experiences a blue or red shift depending on whether the gravitational potential is larger at the point of emission or at the point of reception. This corresponds to a gravitational time 'dilation,' and adds to the time dilation due to velocity.

The metric of general relativity depends on the distribution and motion of masses and is expressed through the gravitational potential, U , and the gravito-magnetic vector potential \mathbf{W} . Then

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta, \quad (7.21)$$

with $dx^\alpha = (cdt, dx, dy, dz)$ and $g_{\alpha\beta} = g_{\alpha\beta}(t, x, y, z, U, \mathbf{W})$.

7.8

IAU Resolutions

In 2000 the IAU adopted the following resolutions specifying the metrics for general relativity for fundamental astronomy.

Resolution B1.3

Definition of Barycentric Celestial Reference System and Geocentric Celestial Reference System

The XXIVth International Astronomical Union General Assembly,

Considering

1. that the Resolution A4 of the XXIst General Assembly (1991) has defined a system of space-time coordinates for (a) the solar system (now called the Barycentric Celestial Reference System, (BCRS) and (b) the Earth (now called the Geocentric Celestial Reference System (GCRS), within the framework of General Relativity,
2. the desire to write the metric tensors both in the BCRS and in the GCRS in a compact and self-consistent form,
3. the fact that considerable work in General Relativity has been done using the harmonic gage that was found to be a useful and simplifying gage for many kinds of applications,

Recommends

1. the choice of harmonic coordinates both for the barycentric and for the geocentric reference systems,
2. writing the time-time component and the space-space component of the barycentric metric $g_{\mu\nu}$ with barycentric coordinates (t, \mathbf{x}) (t = Barycentric Coordinate Time (TCB) with a single scalar potential $w(t, \mathbf{x})$ that generalizes the Newtonian potential and the space-time component with a vector potential $\mathbf{w}^j(t, \mathbf{x})$; as a boundary condition it is assumed that these two potentials vanish far from the solar system,

3. explicitly,

$$g_{00} = -1 + \frac{2w}{c^2} - \frac{2w^2}{c^4},$$

$$g_{0i} = -\frac{4}{c^3}w^i,$$

$$g_{ij} = \delta_{ij} \left(1 + \frac{2}{c^2}w \right),$$

with

$$w(t, \mathbf{x}) = G \int d^3x' \frac{\sigma(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} + \frac{1}{2c^2} G \frac{\partial^2}{\partial t^2} \int d^3x' \sigma(t, \mathbf{x}') |\mathbf{x} - \mathbf{x}'|,$$

$$w^i(t, \mathbf{x}) = G \int d^3x' \frac{\sigma^i(t, \mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|},$$

here, σ and σ^i are the gravitational mass and current densities, respectively,

4. writing the geocentric metric tensor G_{ab} with geocentric coordinates (T, \mathbf{X}) (T = Geocentric Coordinate Time (TCG)) in the same form as the barycentric one but with potentials $W(T, \mathbf{X})$ and $W^a(T, \mathbf{X})$; these geocentric potentials should be split into two parts—potentials W_E and W_E^a arising from the gravitational action of the Earth and external parts W_{ext} and W_{ext}^a due to tidal and inertial effects; the external parts of the metric potentials are assumed to vanish at the geocenter and admit an expansion into positive powers of \mathbf{X} , explicitly,

$$G_{00} = -1 + \frac{2W}{c^2} - \frac{2W^2}{c^4},$$

$$G_{0a} = -\frac{4}{c^3}W^a,$$

$$G_{ab} = \delta_{ab} \left(1 + \frac{2}{c^2}W \right),$$

the potentials W and W^a should be split according to

$$W(T, \mathbf{X}) = W_E(T, \mathbf{X}) + W_{\text{ext}}(T, \mathbf{X}),$$

$$W^a(T, \mathbf{X}) = W_E^a(T, \mathbf{X}) + W_{\text{ext}}^a(T, \mathbf{X}),$$

the Earth's potentials W_E and W_E^a are defined in the same way as w and w^i but with quantities calculated in the GCRS with integrals taken over the whole Earth,

5. using, if accuracy requires, the full post-Newtonian coordinate transformation between the BCRS and the GCRS as induced by the form of the corresponding metric tensors, explicitly, for the kinematically nonrotating GCRS (T = TCG, t = TCB, $r_E^i \equiv x^i - x_E^i(t)$, and a summation from 1 to 3 over equal indices is implied),

$$T = t - \frac{1}{c^2} [A(t) + v_E^i r_E^i] + \frac{1}{c^4} [B(t) + B^i(t) r_E^i + B^{ij}(t) r_E^i r_E^j + C(t, x)] + O(c^{-5}),$$

$$X^a = \delta_{ai} \left[r_E^i + \frac{1}{c^2} \left(\frac{1}{2} v_E^i v_E^j r_E^j + w_{\text{ext}}(x_E) r_E^i + r_E^i a_E^j r_E^j - \frac{1}{2} a_E^i r_E^2 \right) \right] + O(c^{-4}),$$

where

$$\frac{d}{dt} A(t) = \frac{1}{2} v_E^2 + w_{\text{ext}}(x_E),$$

$$\frac{d}{dt} B(t) = -\frac{1}{8} v_E^4 - \frac{3}{2} v_E^2 w_{\text{ext}}(x_E) + 4 v_E^i w_{\text{ext}}^i(x_E) + \frac{1}{2} w_{\text{ext}}^2(x_E),$$

$$B^i(t) = -\frac{1}{2} v_E^2 v_E^i + 4 w_{\text{ext}}^i(x_E) - 3 v_E^i w_{\text{ext}}(x_E),$$

$$B^{ij}(t) = -v_E^i \delta_{aj} Q^a + 2 \frac{\partial}{\partial x^j} w_{\text{ext}}^i(x_E) - v_E^i \frac{\partial}{\partial x^j} w_{\text{ext}}(x_E) + \frac{1}{2} \delta^{ij} \dot{w}_{\text{ext}}(x_E),$$

$$C(t, x) = -\frac{1}{10} r_E^2 (\dot{a}_E^i r_E^i),$$

here x_E^i , v_E^i , and a_E^i are the components of the barycentric position, velocity and acceleration vectors of the Earth, the dot stands for the total derivative with respect to t , and

$$Q^a = \delta_{ai} \left[\frac{\partial}{\partial x_i} w_{\text{ext}}(x_E) - a_E^i \right].$$

The external potentials, w_{ext} and w_{ext}^i , are given by

$$w_{\text{ext}} = \sum_{A \neq E} w_A, \quad w_{\text{ext}}^i = \sum_{A \neq E} w_A^i,$$

where E stands for the Earth and w_A and w_A^i are determined by the expressions for w and with integrals taken over body A only.

Notes

It is to be understood that these expressions for w and w^i give g_{00} correct up to $O(c^{-5})$, g_{0i} up to $O(c^{-5})$, and g_{ij} up to $O(c^{-4})$. The densities σ and σ^i are determined by the components of the energy momentum tensor of the matter composing the solar system bodies as given in the references. Accuracies for G_{ab} in terms of c^{-n} correspond to those of $g_{\mu\nu}$.

The external potentials W_{ext} and W_{ext}^a can be written in the form

$$W_{\text{ext}} = W_{\text{tidal}} + W_{\text{iner}},$$

$$W_{\text{ext}}^a = W_{\text{tidal}}^a + W_{\text{iner}}^a.$$

W_{tidal} generalizes the Newtonian expression for the tidal potential. Post-Newtonian expressions for W_{tidal} and W_{tidal}^a can be found in the references. The potentials W_{iner}^a are inertial contributions that are linear in X^a . The former is determined mainly by the coupling of the Earth's nonsphericity to the external potential. In the kinematically nonrotating Geocentric Celestial Reference System, W_{iner}^a describes the Coriolis force induced mainly by geodetic precession.

Finally, the local gravitational potentials W_E and W_E^a of the Earth are related to the barycentric gravitational potentials w_E and w_E^i by

$$W_E(T, \mathbf{X}) = w_E(t, \mathbf{x}) \left(1 + \frac{2}{c^2} v_E^2 \right) - \frac{4}{c^2} v_E^i w_E^i(t, \mathbf{x}) + O(c^{-4}),$$

$$W_E^a(T, \mathbf{X}) = \delta_{ai} (w_E^i(t, \mathbf{x}) - v_E^i w_E(t, \mathbf{x})) + O(c^{-2}).$$

Resolution B1.4

Post-Newtonian Potential Coefficients

The XXIVth International Astronomical Union General Assembly,

Considering

1. that for many applications in the fields of celestial mechanics and astrometry a suitable parametrization of the metric potentials (or multipole moments) outside the massive solar system bodies in the form of expansions in terms of potential coefficients are extremely useful, and
2. that physically meaningful post-Newtonian potential coefficients can be derived from the literature,

Recommends

1. expansion of the post-Newtonian potential of the Earth in the Geocentric Celestial Reference System (GCRS) outside the Earth in the form

$$W_E(T, \mathbf{X}) = \frac{GM_E}{R} \left[1 + \sum_{l=2}^{\infty} \sum_{m=0}^l \left(\frac{R_E}{R} \right)^l P_{lm}(\cos\theta) (C_{lm}^E(T) \cos m\phi + S_{lm}^E(T) \sin m\phi) \right].$$

here C_{lm}^E and S_{lm}^E are, to sufficient accuracy, equivalent to the post-Newtonian multipole moments introduced by Damour, Soffel and Xu (1991). θ and ϕ are the polar angles corresponding to the spatial coordinates X^a of the GCRS and $R = |\mathbf{X}|$, and

2. expression of the vector potential outside the Earth, leading to the well-known Lense-Thirring effect, in terms of the Earth's total angular momentum vector S_E in the form

$$W_E^a(T, \mathbf{X}) = -\frac{G}{2} \frac{(\mathbf{X} \times \mathbf{S}_E)^a}{R^3}.$$

Resolution B1.5**Extended Relativistic Framework for Time Transformations and Realization of Coordinate Times in the Solar System**

The XXIVth International Astronomical Union General Assembly,

Considering

1. that the Resolution A4 of the XXIst General Assembly, 1991 (Trans IAU, 1992) has defined systems of space-time coordinates for the solar system (Barycentric Reference System) and for the Earth (Geocentric Reference System), within the framework of General Relativity,
2. that Resolution B1.3 entitled 'Definition of Barycentric Celestial Reference System and Geocentric Celestial Reference System' has renamed these systems the Barycentric Celestial Reference System (BCRS) and the Geocentric Celestial Reference System (GCRS), respectively, and has specified a general framework for expressing their metric tensor and defining coordinate transformations at the first post-Newtonian level,
3. that, based on the anticipated performance of atomic clocks, future time and frequency measurements will require practical application of this framework in the BCRS,
4. that theoretical work requiring such expansions has already been performed,

Recommends

that for applications that concern time transformations and realization of coordinate times within the solar system, Resolution B1.3 be applied as follows:

1. the metric tensor be expressed as

$$g_{00} = -\left(1 - \frac{2}{c^2}(w_0(t, \mathbf{x}) + w_L(t, \mathbf{x})) + \frac{2}{c^4}(w_0^2(t, \mathbf{x}) + \Delta(t, \mathbf{x}))\right),$$

$$g_{0i} = -\frac{4}{c^3}w^i(t, \mathbf{x}),$$

$$g_{ij} = \left(1 + \frac{2w_0(t, \mathbf{x})}{c^2}\right)\delta_{ij},$$

where ($t \equiv$ Barycentric Coordinate Time (TCB), \mathbf{x}) are the barycentric coordinates, $w_0 = G \sum_A M_A / r_A$, with the summation carried out over all solar system bodies A, $\mathbf{r}_A = \mathbf{x} - \mathbf{x}_A$, \mathbf{x}_A are the coordinates of the center of mass of body A, $r_A = |\mathbf{r}_A|$, and where w_L contains the expansion in terms of multipole moments (see their definition in the Resolution B1.4 entitled 'Post-Newtonian Potential Coefficients') required for each body. The vector potential $w^i(t, \mathbf{x}) = \sum_A w_A^i(t, \mathbf{x})$, and the function $\Delta(t, \mathbf{x}) = \sum_A \Delta_A(t, \mathbf{x})$ are given in note 2.

2. the relation between TCB and Geocentric Coordinate Time (TCG) can be expressed to sufficient accuracy by

$$\begin{aligned} \text{TCB} - \text{TCG} = c^{-2} & \left[\int_{t_0}^t \left(\frac{v_E^2}{2} + w_{0\text{ext}}(\mathbf{x}_E) \right) dt + v_E^i r_E^i \right] \\ & - c^{-4} \left[\int_{t_0}^t \left(-\frac{1}{8} v_E^4 - \frac{3}{2} v_E^2 w_{0\text{ext}}(\mathbf{x}_E) + 4 v_E^i w_{\text{ext}}^i(\mathbf{x}_E) + \frac{1}{2} w_{0\text{ext}}^2(\mathbf{x}_E) \right) dt \right. \\ & \left. - \left(3 w_{0\text{ext}}(\mathbf{x}_E) + \frac{v_E^2}{2} \right) v_E^i r_E^i \right], \end{aligned}$$

where v_E is the barycentric velocity of the Earth and where the index ‘ext’ refers to summation over all bodies except the Earth.

Notes

1. This formulation will provide an uncertainty not larger than 5×10^{-18} in rate and, for quasi-periodic terms, not larger than 5×10^{-18} in rate amplitude and 0.2 ps in phase amplitude, for locations farther than a few solar radii from the Sun. The same uncertainty also applies to the transformation between TCB and TCG for locations within 50 000 km of the Earth. Uncertainties in the values of astronomical quantities may induce larger errors in the formulas.
2. Within the above-mentioned uncertainties, it is sufficient to express the vector potential $w_A^i(t, \mathbf{x})$ of body A as

$$w_A^i(t, \mathbf{x}) = G \left[\frac{-(\mathbf{r}_A \times \mathbf{S}_A)^i}{2r_A^3} + \frac{M_A v_A^i}{r_A} \right],$$

where \mathbf{S}_A is the total angular momentum of body A and v_A^i are the components of the barycentric coordinate velocity of body A. As for the function $\Delta_A(t, \mathbf{x})$ it is sufficient to express it as

$$\Delta_A(t, \mathbf{x}) = \frac{GM_A}{r_A} \left[-2v_a^2 + \sum_{B \neq A} \frac{GM_B}{r_{BA}} + \frac{1}{2} \left(\frac{(r_A^k v_A^k)^2}{r_A^2} + r_A^k a_A^k \right) \right] + \frac{2Gv_A^k (\mathbf{r}_A \times \mathbf{S}_A)^k}{r_A^3},$$

where $\mathbf{r}_{BA} = |\mathbf{x}_B - \mathbf{x}_A|$ and a_A^k is the barycentric coordinate acceleration of body A. In these formulas, the terms in \mathbf{S}_A are needed only for Jupiter ($S \approx 6.9 \times 10^{38} \text{ m}^2 \text{ s}^{-1} \text{ kg}$) and Saturn ($S \approx 1.4 \times 10^{38} \text{ m}^2 \text{ s}^{-1} \text{ kg}$), in the immediate vicinity of these planets.

3. Because the present Recommendation provides an extension of the IAU 1991 recommendations valid at the full first post-Newtonian level, the constants L_C and L_B that were introduced in the IAU 1991 recommendations should be defined as $\langle \text{TCG} / \text{TCB} \rangle = 1 - L_C$ and $\langle \text{TT} / \text{TCB} \rangle = 1 - L_B$, where TT refers to Terrestrial Time and $\langle \rangle$ refers to a sufficiently long average taken at the geocenter. The most recent estimate of L_C is (Irwin and Fukushima, 1999)

$$L_C = 1.48082686741 \times 10^{-8} \pm 2 \times 10^{-17}.$$

From the Resolution B1.9 on ‘Redefinition of Terrestrial Time TT ’, one infers that $L_B = 1.55051976772 \times 10^{-8} \pm 2 \times 10^{-17}$ by using the relation $1 - L_B = (1 - L_C)(1 - L_G)$. L_G is defined in Resolution B1.9.

Because no unambiguous definition may be provided for L_B and L_C , these constants should not be used in formulating time transformations when it would require knowing their value with an uncertainty of order 1×10^{-16} or less.

4. If TCB–TCG is computed using planetary ephemerides which are expressed in terms of a time argument (noted T_{eph}) which is close to Barycentric Dynamical Time (TDB), rather than in terms of TCB, the first integral in Recommendation 2 above may be computed as

$$\int_{t_0}^t \left(\frac{v_E^2}{2} + w_{0\text{ext}}(\mathbf{x}_E) \right) dt = \left[\int_{T_{\text{epho}}}^{T_{\text{eph}}} \left(\frac{v_E^2}{2} + w_{0\text{ext}}(\mathbf{x}_E) \right) dt \right] / (1 - L_B).$$

7.9

Time Scales

7.9.1

International Atomic Time

International Atomic Time (TAI) was established originally from the new cesium atomic clocks operating in laboratories on the surface of the Earth. There was no consideration of relativistic effects in the first applications of the new atomic standards in establishing an atomic time scale. However, as accuracies improved and atomic time standards were located at many different sites on Earth and in orbit around the Earth, corrections based on differences in potentials had to be introduced for comparisons between the atomic standards (See Chapter 12).

7.9.2

Dynamical Time Scales

With the adoption of the resolutions defining the Barycentric and Geocentric Celestial Reference Systems, two time scales, Barycentric Coordinate Time (TCB) and Geocentric Coordinate Time (TCG), were also introduced and the relationships between these time scales and Terrestrial Time (TT) were clarified. Barycentric Dynamical Time (TDB) was defined in 2006 as a linear scaling of TCB to have the approximate rate of TT. These time scales are specifically discussed in Chapter 8.

7.10

Relativistic Effects in Time Transfer

Time transfer can be considered in two different reference frames, either a geocentric, Earth-fixed, rotating frame or a geocentric, nonrotating, local inertial frame. Clearly, clock comparisons and synchronization must be made in a common

coordinate system. For Earth-based purposes and for orbiting satellites, the rotating or nonrotating geocentric coordinate system should be used. Since they have the same coordinate time, either may be used. In practice, for ground-based clocks the rotating system is preferable, while for orbiting clocks the nonrotating system is preferred. The equations given will include both frames and consider clocks up to geosynchronous altitudes. (See Chapter 15 for details on Time Transfer.)

References

- Damour, T., Soffel, M. and Xu, C. (1991) General-relativistic celestial mechanics. I. Method and definition of reference systems. *Phys. Rev. D*, **43**, 3273–307.
- Irwin, A. and Fukushima, T. (1999) A numerical time ephemeris of the Earth. *Astron. Astroph.*, **348**, 642–52.
- Kovalevsky, J. and Seidelmann, P.K. (2004) *Fundamentals of Astrometry*, Cambridge University Press, Cambridge.
- Minkowski, H. (1908) Space and time, in *The Principle of Relativity* (eds H.A. Lorentz, A. Einstein, H. Minkowski and H. Weyl, translated by W. Perrett and G.B. Jeffery), Methuen & Co., London, 1923, pp. 75–96.
- Trans. Int. Astron. Union (1992) *Proc. 21st General Assembly Buenos Aires, 1991*, Vol. XXI B (ed. J. Bergeron), Kluwer Academic Publishers.
- Trans. Int. Astron. Union (2001) *Proc. 24th General Assembly Manchester, 2000*, Vol. XXIV B (ed. H. Rickman), Astron. Soc. of the Pacific.

8

Dynamical and Coordinate Time Scales

8.1

Replacing Ephemeris Time

With the recognition of the problems with Ephemeris Time (ET) and the acknowledged need for its replacement, new time scales needed to be defined. Originally conceived as a uniform time scale for use in celestial dynamics and possibly in other practical applications, it was becoming clear in the 1960s and 1970s that atomic time provided a much more accessible time scale for practical use. For dynamical applications, ET was the independent argument of the equations of motion for the ephemerides used in its definition, but because its definition was dependent on out-dated astronomical constants and Newtonian mechanics it could not be considered the independent argument in the fundamental equations of motion. So, when changes in the celestial reference system were to be considered at the IAU General Assembly in Grenoble, France, in 1976, the time was appropriate to think about the possibility of improved definitions of time scales. It was clearly desirable to make new dynamical time scales continuous with ephemeris time within the accuracy of the previous time scale. Also, any new time scales needed to be consistent with the theory of relativity.

In the context of this chapter, dynamical time is understood as the time-like argument of dynamical theories. Dynamical time scales represent the independent variable of the equations of motion of solar system bodies and depend on the reference system being used. The equations of motion can be referred to the barycenter of the solar system or to the geocenter. Each system would require its appropriate coordinate time scale. The process of arriving at the appropriate definitions involved some confusion, mistakes, misunderstandings, and corrections. Some unresolved issues remain today.

Some choices for possible new dynamical time scales were (i) to improve the definition of ET, (ii) to adopt International Atomic Time (TAI), or (iii) to seek a new definition for time scales or time-like arguments. The IAU considered a number of factors in approaching the problem. Coordinated Universal Time (UTC), the adopted time scale, was based on atomic time but subject to one-second discontinuities to maintain its relationship to the Earth's rotation. However, a uniform time scale was needed to compute ephemerides in the past as well as the

future, and atomic time was not available before 1955. TAI (which stands for the French term *Temps Atomique International*), the conventional realization of atomic time, was (and remains) a time scale based on contributed clock data from around the world and subject to the uncertainties inherent in combining many different time standards. (See Chapter 12.) The known systematic discrepancies in the ephemerides of solar system objects made the realization of a dynamic time scale using solar system ephemerides problematical. The defining relationships between time scales needed to account for the effects of relativity, and there was a question of the equivalence of time based on quantum physics and time based on dynamics.

With these considerations, it was decided to recommend new time-like arguments for the independent variable for solar system body theories. The definitions would distinguish between coordinate and proper time scales, be related to TAI at some epoch, and be continuous with ephemeris time. It was possible, then, to have either a barycentric, or a geocentric, dynamical time scale. In practice, with the need to describe the motions of bodies in both geocentric and barycentric reference systems, both types of time scales were required. The details of the relationship between them would depend on the specifications of the relativistic theory.

8.2

Terrestrial Dynamical Time (TDT) and Barycentric Dynamical Time (TDB)

Recommendation 5, Time Scale for Dynamical Theories and Ephemerides, adopted by the IAU General Assembly of 1976 in Grenoble, France reads:

‘It is recommended that

- (a) at the instant 1977 January 01^d 00^h 00^m 00^s TAI, the value of the new time scale for the apparent geocentric ephemerides will be 1977 January 1.^d0003725 exactly (1^d 00^h 00^m 32.^s184);
- (b) the unit of this time scale will be a day of 86 400 SI seconds at mean sea level;
- (c) the time scales for equations of motion referred to the barycenter of the solar system will be such that there will be only periodic variations between these time scales and that for the apparent geocentric ephemerides;
- (d) no time-step will be introduced in International Atomic Time.’

The recommendations also included notes of explanation that were not part of the recommendation itself. Items of interest included in the notes are the following:

1. The time-like arguments of dynamical theories and ephemerides are referred to as dynamical time scales. While it is possible and desirable to base the unit of dynamical time scales on the SI second, it is necessary to recognize that in relativistic theories there will be periodic variations between the unit of time

for an apparent geocentric ephemeris and the unit of the corresponding time scale of the equation of motion, which may, for example, be referred to the center of mass of the solar system. (In the terminology of the theory of general relativity such time scales may be considered to be proper time and coordinate time, respectively.) The time scales for an apparent geocentric ephemeris and for the equations of motion will be related by a transformation that depends on the system being modeled and on the theory being used. The arbitrary constants in the transformation may be chosen so that the time scales have only periodic variations with respect to each other. Thus it is sufficient to specify the basis of a unique time scale to be used for new, precise, apparent geocentric ephemerides. The dynamical time scale for apparent geocentric ephemerides of recommendation 5 (a) and 5 (b) is a unique time scale independent of theories, while the dynamical time scales referred to the barycenter of the solar system are a family of time scales resulting from the transformation of various theories and metrics of relativistic theories.

2. This recommendation specifies a particular dynamical time scale for apparent geocentric ephemerides that is effectively equal to TAI +32.184s. (There are formal differences arising from random, and possibly systematic, errors in the length of the TAI second and the method of forming TAI, but the accumulated effect of such errors is likely to be insignificant for astronomical purposes over long periods of time.) The scale is specified with respect to TAI in order to take advantage of the direct availability of UTC (which is based on the SI second and is simply related to TAI) and to provide continuity with the current values and practice in the use of Ephemeris Time. Continuity is achieved since the chosen offset between the new scale and TAI is the current estimate of the difference between ET and TAI and the SI second was defined so as to make it equal to the ephemeris second within the error of measurement. It will be possible to use most available ephemerides as if the arguments were on the new scale. Before 1955, when atomic time was not available, the determinations of ET can be considered to refer to the new time scale. The offset has been expressed in the recommendation as an exact decimal fraction of a day since the arguments of theories and ephemerides are normally expressed in days.
3. In view of the desirability of maintaining continuity of TAI and of avoiding the confusion that would arise if it were to be redefined retrospectively, no step in TAI is proposed. Although the recommendation is in terms of TAI, in practice astronomers will use UTC and convert directly to the dynamical time scale.
4. The terminology and notation for dynamical time scales require further consideration in due course.
5. Recognizing that the TAI second differed from the SI second between 1969 and the present by $(10 \pm 2) \times 10^{-13}$, a step will be introduced in the scale interval of TAI. Therefore, the epoch of the dynamical time scale for apparent geocentric ephemerides was adjusted to 1977 at Grenoble (*Trans. Int. Astron. Union*, 1977).'

In 1979 at the IAU General Assembly in Montreal, although there was discussion about the word ‘dynamical’ and possible ambiguities or misunderstandings, the names of Terrestrial Dynamical Time (TDT) and Barycentric Dynamical Time (TDB) were adopted as the names of the two time scales (*Trans. Int. Astron. Union*, 1980). As a result of the IAU recommendations of 1976 and 1979, TDT was defined to provide continuity with ET, with the chosen offset thought to be an accurate value of the difference in TAI and ET at the time of the introduction of TDT. The extrapolation of TDT backward, prior to the availability of TAI, must be made by the use of ephemerides of the solar system as was done for ET. Barycentric Dynamical Time (TDB) was defined to be the independent variable of the equations of motion for solar system bodies with respect to the solar system barycenter. In practice, TDB was to be determined from TDT by a mathematical expression which depends upon the gravitational theory being used, the astronomical constants, and the positions and motions of the solar system bodies. Mathematically these time scales could be realized for any TDB epoch, t_{TDB} , with the expressions in Eq. (8.1) (Kaplan, 1981). More accurate expressions are available for TDB, for example, Moyer (1981), Fairhead *et al.* (1988), and Hirayama *et al.* (1987). These two time-like arguments were to be introduced in use for practical applications in astronomy in the beginning of 1984.

$$\begin{aligned}
 \text{TDT} &= \text{TAI} + 32.184 \text{ s}, \\
 \text{TDB} &\approx \text{TDT} + 0.001658 \text{ s} \sin(g + 0.0167 \sin g) \\
 &\quad + \text{lunar and planetary terms of order } 10^{-5} \text{ s} \\
 &\quad + \text{daily terms of order } 10^{-6} \text{ s} \\
 g &= \text{mean anomaly of the Earth in its orbit} \\
 &= (357.528^\circ + 35\,999.050^\circ T) \times \frac{2\pi}{360^\circ}, \\
 T &= \frac{(t_{\text{TDB}} - 2\,451\,545.0)}{36525}.
 \end{aligned} \tag{8.1}$$

8.3

Problems with TDT and TDB

Although the 1976 and 1979 resolutions were not ambiguous, the notes and the use of the word ‘dynamical’ to designate the time scales led to different interpretations and confusion. This issue was discussed in a paper by Guinot and Seidelmann (1988). They noted that TDT is defined by its origin with respect to TAI and its unit, which at any instant is equal to the day of 86 400 seconds of SI, at sea level. Thus, TDT is an ideal form of atomic time (of which TAI with an offset is a practical realization), but it has no relationship with the motion of solar system bodies. Thus, the use of the word ‘dynamical’ in the name was inappropriate, if we understand that the use of the word ‘dynamical’ implies a relationship to

dynamical theory, as ‘atomic’ time would imply a relationship with atomic theory. The notes to the Grenoble recommendation used the word dynamical both as a time-like argument for dynamical theories and ephemerides and as a unique time scale independent of theories (i.e., as an ideal atomic time).

There was also confusion about whether the unit of time of 86 400 SI seconds was specified only for the epoch in 1977, or as being constant for all times. The realization of a unit of time in a dynamical system could be expected to vary because of incomplete modeling of the dynamics and possible errors in the model used in preparing the ephemerides. There were also serious problems with the definition of TDB. It specified that there should only be periodic variations between the two time scales, but it did not specify over what period of time there should only be periodic differences and the nature of periodic terms which would be acceptable.

Guinot and Seidelmann also discussed the facts that there were different interpretations and understandings of relativistic theories and that there were many meanings of the word ‘dynamical.’ They noted the widely different requirements for time scales, and that it was necessary to distinguish between (i) the need for time scales for the apparent motions of celestial bodies in a space and time reference system for which realizations are available and (ii) the guidance to theoreticians concerning the time-like arguments for dynamical theories. Guinot and Seidelmann suggested that for future recommendations regarding time scales the IAU should consider the following characteristics for an ideal time:

- ‘(a) This ideal time must have realizations (time scales) available to all terrestrial observers, with a good precision, and such that the deviation between the ideal time and the corresponding realized time scales remains as small as possible and negligible for most of the applications.
- (b) The definition must be unambiguous and must contain all the information which is needed so that the ideal time can be considered to be the argument of the ephemerides.’

They then went on to propose a recommendation designed to improve on the IAU recommendations of 1976 and 1979.

8.4 New Reference System

In the 1980s it became apparent that there was a need to reconsider the celestial reference system. Radio frequency observations by interferometers were achieving accuracies much better than the optical positional observations, and there were plans for astrometric space missions that would also be much more accurate than the historical fundamental optical catalogs. Accuracies were changing from tenths of an arcsecond to milliarcseconds. An IAU Working Group on Reference Systems was established with four subgroups: one on reference frames and their origin, another on time scales, a third on astronomical constants, and a fourth on the

theory of nutation. IAU Colloquium 127 on Reference Systems (Hughes, Smith and Kaplan, 1991) produced nine recommendations that defined space-time coordinates within the framework of general relativity for the barycenter of any ensemble of masses. The recommendations were adopted by the IAU in Buenos Aires in 1991 (*Trans. Int. Astron. Union*, 1992).

These recommendations explicitly introduced the general theory of relativity as the theoretical background for the definition of space-time reference frames. It was recommended that the celestial reference frames with origins at the solar system barycenter and the center of mass of the Earth should show no global rotation with respect to a set of distant extragalactic objects, and work was requested to be initiated for the observational basis for the reference frame. The IAU recommended that the equator of this new conventional celestial barycentric reference frame should be as near as possible to the mean equator of J2000.0 and that the origin on this plane be as near as possible to the dynamical equinox of J2000.0, so that there would be no discontinuity of the new reference frame with the FK5. After the initial alignment of the new frame, no further rotation was to be applied to match the frame with a better determination of the mean equator and dynamical equinox of J2000.0. In related issues the IAU Working Group on Astronomical Standards recommended the use of observed celestial pole offsets in cases where accuracies better than 2 milliarcseconds (mas) were required.

The resolutions began a series of activities that led to the International Celestial Reference System (ICRS), International Celestial Reference Frame (ICRF), Hipparcos reference star catalog, fainter reference star catalogs consistent with the ICRF, the IAU 2000A precession-nutation model, IAU 2006 precession theory, and continuing reference system improvements.

8.5

New Time Scales

In time-related issues, the IAU recommended that the time coordinates for the celestial barycentric reference system should be derived from a time scale realized by atomic clocks operating on the Earth. The basic physical units of space time of all coordinate systems centered at the barycenters of ensembles of masses were to be chosen so that they would be consistent with the proper unit of time, the SI second, and for proper length, the SI meter, related to the SI second by the speed of light in vacuum. The SI second is specified as the duration of radiation periods of cesium, but it contains no specification of location, gravitational potential, or state of motion of the observer. Therefore, the SI second can be the unit of proper time for any location.

The readings of the coordinate times were set to be 1977 January 1, 0^h 0^m 32^s.184 TAI exactly at the geocenter. The coordinate times corresponding to the coordinate systems, which have their spatial origins respectively at the center of mass of the Earth and at the solar system barycenter, were designated as Geocentric Coordinate Time (TCG) and Barycentric Coordinate Time (TCB).

The coordinate time scales TCB, TDB, TCG, and TT are defined for any space time event. Coordinate times cannot be measured; they can only be computed from readings of real clocks. The computation must be based on the theoretical relation between proper and coordinate time from the principles of general relativity. The relation involves the model of the solar system, the motions of the observer and the massive bodies, and the mass parameters. This model is given in the BCRS for TCB and in the GCRS for TCG, respectively. TDB and TT are related to TCB and TCG, respectively, by linear expressions.

8.5.1

Coordinate Time

In the framework of general relativity, coordinate time is one of the coordinates in a four-dimensional space time system. In contrast, proper time is time measured by a clock between events that occur at the same place as the clock. Since we are on the Earth, geocentric reference frames and time scales have a special importance, and so it is particularly relevant to establish the relationships among time scales within that context.

The center of mass of the Earth can be considered as freely falling in space. Despite its irregular motion within the Earth, amounting to a centimeter or so, we can consider it as a well-defined point and neglect the internal forces and motions that affect it. If we have a clock at rest on the surface of the Earth experiencing a potential U and measuring a proper time interval $d\tau$, then we can relate that time interval to a geocentric coordinate time interval dt neglecting terms in c^{-4} by using the expression

$$\frac{d\tau}{dt} = 1 - \frac{U}{c^2}, \quad (8.2)$$

where c is the speed of light and U is the total potential composed of the gravitational potential U_G and the axifugal force due to the Earth's rotation as experienced at the clock's location (Audoin and Guinot, 2001). The latter component is given by

$$U_R = \frac{1}{2}\omega^2(x^2 + y^2), \quad (8.3)$$

for a clock with geocentric coordinates x, y, z , (z being the distance above the plane of the equator) and the Earth's rotational speed being given by ω .

8.5.2

Terrestrial Time

The coordinate time defined at the surface of the Earth is called Terrestrial Time (TT) and differs from TCG uniquely by a constant rate. The unit of time of TT is the second of the International System of Units (SI) as realized on the geoid and defined 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 cesium 133 atom.’

This definition refers to the atom in its ground state undisturbed and at a temperature of 0° Kelvin. Atomic cesium clocks realize the SI second as its proper time, which, at the geoid, is equal to the coordinate time TT.

Terrestrial Time (TT) was defined to be an ideal time-like argument. The unit of measurement of TT agrees with the SI second on the geoid, and, at the instant 1977 January 1, 0^h 0^m 0^s TAI exactly, TT had the reading 1977 January 1, 0^h 0^m 32^s.184 exactly. TT was defined as compatible with the definition of TAI given by CCDS (1980), but the definition was more accurate, because the form of the metric was given. Thus, TT has only a small potential drift with respect to the statistical time scale TAI. TT was to be used as the time reference for apparent geocentric ephemerides, and to differ from TCG by a constant rate. TT is a coordinate time scale, but not a dynamical time scale, and serves as the bridge from TAI to the coordinate time scales. TAI is the practical realization of TT, which is, in practice, derived by the BIPM (Bureau International des Poids et Mesures) by comparing the readings of many atomic clocks, reduced to the geoid. The two time scales are related by

$$TT = TAI + 32.184s. \quad (8.4)$$

The constant 32.184s is the result of the fact that, when TAI was initiated in 1958, it was set equal to the epoch of Universal Time at that time instead of Ephemeris Time because Ephemeris Time was so poorly determined. However it was later decided that it would be preferable to have TAI be continuous with ET. Consequently the constant 32.184s was adopted as the best estimate of the accumulated difference between TAI and ET as of 1 January 1977.

8.5.3

Geocentric Coordinate Time

TCG is the time coordinate for the four-dimensional geocentric coordinate system, and in terms of the theory of relativity, differs slightly in rate from TT. It was set to coincide with TT on 1977 January 1, 0^d 0^h 32^s.184 (JD 2442144.5003725). TT and TCG differ only by a scaling factor so that

$$TCG = TT + L_G \times (JD - 2\,443\,144.5) \times 86\,400s. \quad (8.5)$$

$L_G = 6.969\,290\,134 \times 10^{-10}$ exactly, and is specified as a defining constant that will not change with future improved Earth models. This overcomes difficulties due to temporal changes of the geoid and the intricacy of its definition.

8.5.4

Barycentric Coordinate Time

TCB is the time coordinate for the four-dimensional barycentric coordinate system of the solar system, differing both in secular and periodic effects from TT and

TCG, according to the relativistic metric being used. A particularly important reference system in astronomy is that centered at the barycenter of the solar system. This point is chosen in such a way that its motion in the Galaxy is linear if one neglects the gravitational potential of the Galaxy. One of the problems of astrometry is to refer observations to this barycentric system whose associated coordinate time scale is called TCB. A complication in computing TCB for any event in the solar system is that the time-varying gravitational potentials, which must be accounted for, depend on the positions of the planets.

If TCB and \mathbf{x}_0 are the barycentric coordinates of an event in a barycentric reference system the potential is described by the sum of potentials produced by N significant solar system bodies with assumed point-like masses M_i and positions \mathbf{x}_i :

$$U_0 = G \sum_{i=1}^N \frac{M_i}{|\mathbf{x}_i - \mathbf{x}_0|}. \quad (8.6)$$

If the event refers to the center of mass of the Earth, the potential is given by the summation above, but would not include the gravitational potential of the Earth. This potential can be designated U_0^{ext} . The relationship between TCB and TCG can be given by the full four-dimensional transformation:

$$\text{TCB} - \text{TCG} = c^{-2} \left[\int_{t_0}^t \left(\frac{v_e^2}{2} + U_0^{\text{ext}}(\mathbf{x}_e) \right) dt + \mathbf{v}_e \cdot (\mathbf{x} - \mathbf{x}_e) \right] + O(c^{-4}), \quad (8.7)$$

where c is the speed of light, \mathbf{x}_e and \mathbf{v}_e denote the barycentric position and velocity of the Earth's center of mass, and \mathbf{x} is the barycentric position of the observer. In the integral, $t = \text{TCB}$ and t_0 is chosen to agree with the epoch of Terrestrial Time (TT). $\mathbf{x} - \mathbf{x}_e$ would be zero for the geocenter, as would be the normal case. The additional terms would be of order c^{-4} . This expression is ephemeris dependent, since it depends on \mathbf{x}_e and \mathbf{v}_e . Consequently, there is a 'time ephemeris' associated with every spatial ephemeris of solar system bodies expressed in TCB. It would be expected that spatial ephemerides will be accompanied by 'time ephemerides' for the user.

As an approximation to TCB–TCG in seconds, one might use (McCarthy and Petit, 2004)

$$\text{TCB} - \text{TCG} = \frac{L_C \times (\text{TT} - \text{TT}_0) + P(\text{TT}) - (\text{TT}_0)}{1 - L_B} + c^{-2} \mathbf{v}_e \cdot (\mathbf{x} - \mathbf{x}_e). \quad (8.8)$$

Additional terms in c^{-4} can be found in recommendation B1-4 of the IAU (Trans IAU, 2001). This expression has a linear term which represents the difference of their mean rates.

$$(\text{TCB} - \text{TCG})_{\text{secular}} = L_C (\text{JD} - 2\,443\,144.5) \times 86\,400\text{s}, \quad (8.9)$$

where the best value at present of $L_C = 1.48082686741 \times 10^{-8}$ with an uncertainty of 10^{-17} . In addition, there is a nonlinear variation described by a number of periodic terms depending on the various periods present in the motion of planets. They are discussed in Fukushima (1995), who finds that there are 515 terms which

are greater in amplitude than 0.1 ns. The most important terms in TCB – TCG are, in seconds,

$$0.001\,658 \sin g + 0.000\,014 \sin 2g, \quad (8.10)$$

where $g = 357.53^\circ + 0.985\,003^\circ (\text{JD} - 2\,451\,545.0)$ represents essentially the mean anomaly of the Earth's orbit. The epoch of TCB was set to be equal to TT and TCG at 1977 January 1, $0^{\text{d}}\,0^{\text{h}}\,32^{\text{s}}.184$ (JD 2443144.5003725).

8.5.5

Barycentric Ephemeris Time

Through the series of revisions of time scales, changes in names of the time scales, and further clarifications and improvements in the applications of relativity to the time scales, the Jet Propulsion Laboratory followed an independent but somewhat parallel path. The basic calculations of the barycentric ephemerides were performed using equations of motion including the needed relativistic terms and using an independent variable that was a relativistic time scale for the solar system barycenter. Barycentric Ephemeris Time (T_{eph}) is a coordinate time related to TCB by an offset and a scale factor. Ephemerides based upon the coordinate time T_{eph} are automatically adjusted in the ephemeris creation process so that the rate of T_{eph} has no overall difference from the rate of Terrestrial Time (TT) (Standish, 1998) and also no overall significant difference from the rate of the redefined Barycentric Dynamical Time (TDB). For this reason, space coordinates obtained from the ephemerides are consistent with TDB. The difference between T_{eph} and TDB is that T_{eph} depends on the ephemeris used while TDB has been defined by a particular relationship with respect to TCB. (See Section 8.5.6 below.)

8.5.6

TDB Redefined

In 1991 the continued use of Barycentric Dynamical Time (TDB) had been authorized, specifically where discontinuity with previous work was deemed undesirable. The convenience of T_{eph} for use in operational ephemerides and the problems with the original definition of TDB prompted the IAU in 2006 to redefine TDB. The conventionally accepted solar system ephemerides produced by the Jet Propulsion Laboratory had continued to be given with T_{eph} . The national almanac offices labeled their stellar position tables with TDB (Astronomical Almanac). Hence, the IAU at the 2006 General Assembly in Prague decided to redefine TDB for appropriate applications with the following recommendation.

In situations calling for the use of a coordinate time scale that is linearly related to Barycentric Coordinate Time (TCB) and remains close to Terrestrial Time (TT) at the geocenter for an extended time span, TDB is defined as the following linear transformation of TCB:

$$\text{TDB} = \text{TCB} - L_B \times (\text{JD}_{\text{TCB}} - T_0) \times 86\,400 + \text{TDB}_0 \quad (8.11)$$

where $T_0 = 2\,443\,144.500\,372\,5$, and $L_B = 1.550\,519\,768 \times 10^{-8}$ and $\text{TDB}_0 = -6.55 \times 10^{-5}$ s are defining constants.

There were the following notes to the resolution (*Trans. Int. Astron. Union*, 2007):

- ‘1. JD_{TCB} is the TCB Julian date. Its value is $T_0 = 2\,443\,144.500\,372\,5$ for the event 1977 January 1 00h 00m 00s TAI at the geocenter, and it increases by one for each 86 400 s of TCB.
2. The fixed value that this definition assigns to L_B is a current estimate of $L_C + L_G - L_C \times L_G$, where L_G is given in IAU Resolution B1.9 (2000) and L_C has been determined (Irwin and Fukushima, 1999,) using the JPL ephemeris DE405. When using the JPL Planetary Ephemeris DE405, the defining L_B value effectively eliminates a linear drift between TDB and TT, evaluated at the geocenter. When realizing TCB using other ephemerides, the difference between TDB and TT, evaluated at the geocenter, may include some linear drift, not expected to exceed 1 ns per year.
3. The difference between TDB and TT, evaluated at the surface of the Earth, remains under 2 ms for several millennia around the present epoch.
4. The independent time argument of the JPL ephemeris DE405, which is called T_{eph} (Standish, 1998), is for practical purposes the same as TDB defined in this Resolution.
5. The constant term TDB_0 is chosen to provide reasonable consistency with the widely used TDB–TT formula of Fairhead and Bretagnon (1990).
n.b. The presence of TDB_0 means that TDB is not synchronized with TT, TCG and TCB at 1977 Jan 1.0 TAI at the geocenter.
6. For solar system ephemerides development the use of TCB is encouraged.’

Although the relationships between the dynamical time scales were specified by the IAU, the relationships between the corresponding spatial units were not specified. This was left for future consideration by the new IAU Commission 52, Relativity for Fundamental Astronomy. Scaling factors can be required to relate quantities compatible with the time scales (Klioner, 2008).

There are currently two opinions concerning TDB. From Klioner (2008), TT is a scaled version of TCG, and TDB is a scaled version of TCB. There is no physical significance to either scaling; they are both for the convenience of making the difference between the proper time of an observer and the coordinate times evaluated along his trajectory as small as possible. The relationships between proper and coordinate times, and between two coordinate times, are independent of the units. The SI second is generally used as the unit of proper time. If the theoretical formulas are used to relate the proper time to the coordinate time, with the proper time units of the SI second, the units of time of the coordinate time would also be the SI second. Thus, TDB units would be the SI second.

The other opinion is that TT, TCG, TCB, TAI, and UTC are all SI units. If a clock that counted 9192631770 cesium transitions in one second were placed on the geoid, it would keep TT, TAI, and UTC. If placed at the center of the Earth, shielded from

the Earth's potential, it would keep TCG. If placed at the solar system barycenter, shielded from the solar potential, it would keep TCB. However, for TDB the clock must count a different number of transitions per second. Thus, the unit of time of TDB is not the SI second. The TDB is counting a different time unit so it can keep up with the SI clocks on Earth. Therefore, there is a 'TDB unit' of time.

8.6

ΔT and Ephemeris Time Revised

The development of coordinate time scales further refined the original concept of Ephemeris Time (Chapter 6) to meet the growing needs for improved accuracy. In the recommendations that implemented the dynamical time scales no mention was made of officially changing the definition of ΔT . It was originally defined as $\Delta T = ET - UT$, and, strictly speaking, that remains the definition for times prior to 1983. From 1984–2000, after TDT was introduced as a replacement for ET in 1984, it was assumed that $\Delta T = TDT - UT$. Similarly, following the introduction of TT in 2001, ΔT became $TT - UT$. The existence of atomic time since 1956 together with expression, $\Delta T = TDT - UT$, and Equation 8.1 also allowed us to estimate ΔT for the years following the introduction of atomic time by $\Delta T = TAI + 32.184s - UT$. Although TAI did not formally exist before 1971, it was considered to be an extension of the Bureau International de l'Heure atomic time scale that had been continuous back to 1955 (Nelson *et al.*, 2001).

Having defined the new set of dynamical and coordinate time scales and resolved the various problems that had existed, there remained the problem that there was no atomic time scale prior to 1955. New numerically integrated ephemerides replaced the theories of Newcomb, and they did not have an expression for the mean longitude of the Sun. Hence, while ET is to be used prior to 1955, there has not been a formal revised definition of Ephemeris Time, as it is currently being used, for the reduction of observations and the determinations of ΔT prior to 1955. A definition for Ephemeris Time Revised (ET_R) that matches what is being used in practice was proposed by Guinot and Seidelmann (1988):

- (a) Ephemeris Time Revised (ET_R) is reckoned from the instant 1958 January 0^h TAI, at which time ET_R has the value 1958 January 0^h 0^m 32^s.184.
- (b) The unit of time of ET_R is the SI second in its present definition (atomic second).

This would make Ephemeris Time continuous with Terrestrial Time prior to 1955 to the accuracies involved. It would also be continuous with TDT, but it would not be continuous with TCB or TCG, which differ secularly from TT. No formal action has been taken defining Ephemeris Time Revised, but the practice for determining a dynamical time and values of ΔT prior to 1955 has basically followed the definition above.

Values of ΔT are given annually in *The Astronomical Almanac* from 1620 to the present. Since UTC has steps due to leap seconds, the values of TAI–UTC are also

given from 1972 to the present, and extrapolated values are given for a few years into the future. ΔT values are also available from the International Earth Rotation and Reference System Service (IERS) and United States Naval Observatory (USNO) web sites.

8.7

Relationships Among Coordinate Time Scales

From the sections above it is clear that the need for a practical replacement for Ephemeris Time that also accounted for the relativistic aspects of timekeeping resulted in a series of related conceptual time scales. These scales can all be related mathematically. Furthermore they can be realized essentially through TT and its practical realization TAI. The details of the formation of TAI and its relation to other time scales are discussed in the chapters following.

The dynamical time scales are related to TT by mathematical expressions. Barycentric Dynamical Time (TDB) has periodic differences from TT, but has a rate very close to the rate of TT. These expressions are illustrated in Figure 8.1 along

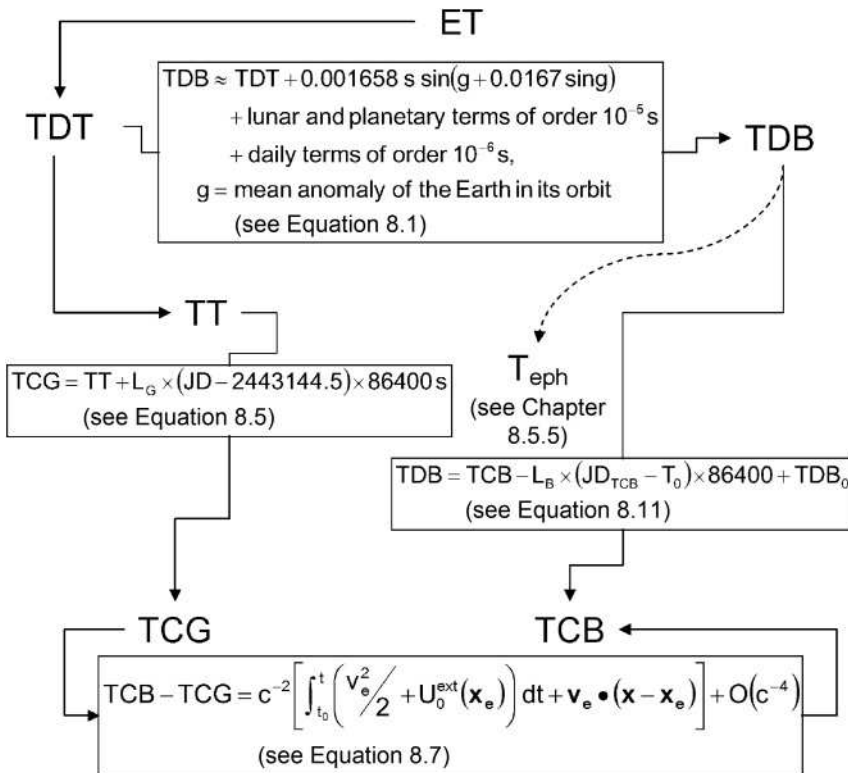


Figure 8.1 Relationships among Ephemeris Time, Terrestrial and Barycentric Dynamical Times, Terrestrial Time, and Terrestrial and Geocentric Coordinate Times.

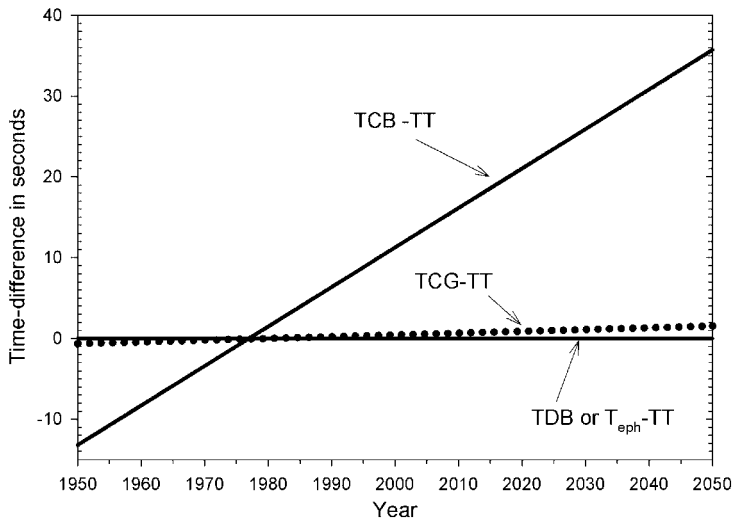


Figure 8.2 Differences in readings of various time scales with respect to TT, TDT, and ET equivalent to TT.

with the expressions relating the coordinate and dynamical time scales. It shows the constant rate difference between Geocentric Coordinate Time (TCG) and TT, and both rate and periodic differences between Barycentric Coordinate Time (TCB) and TCG. The magnitudes of the differences between these different time scales are shown in Figure 8.2, which was originated by Seidelmann and Fukushima (1992).

References

- Audoin, C. and Guinot, B. (2001) *The Measurement of Time*, Cambridge University Press, Cambridge.
- CCDS (1980) *BIPM Com. Cons. Def. Seconde*, **9**, 515.
- Fairhead, L. and Bretagnon, P. (1990) An analytical formula for the time transformation TB-TT. *Astron. Astrophys.*, **229**, 240.
- Fairhead, L., Bretagnon, P. and Lestrade, J.F. (1988) The time transformation TDB-TDT: an analytical formula and related problem of convention, in *The Earth Rotation and Reference Frames for Geodesy and Geodynamics* (eds A.K. Babcock and G.A. Wilkins) Kluwer, Dordrecht, p. 419.
- Fukushima, T. (1995) Time Ephemeris. *Astron. Astrophys.*, **294**, 895–906.
- Guinot, B. and Seidelmann, P.K. (1988) Time Scales: their history, definition, and interpretation. *Astron. Astrophys.*, **194**, 304–8.
- Hirayama, Th., Kinoshita, H., Fujimoto, M.-K. and Fukushima, T. (1987) Analytical expression of TDB-TDT. *Proc. IAG Symposium, IUGG XIX General Assembly, Vancouver, Aug 10–13*, p. 91.
- Hughes, J.A., Smith, C.A. and Kaplan, G.H. (1991) Recommendation of the working group on reference systems. *IAU Proc. 127th Colloquium, Reference Systems, U.S. Naval Observatory, Washington D.C.*, pp. 408–41.
- Irwin, A.W. and Fukushima, T. (1999) A numerical time ephemeris of the Earth. *Astron. Astrophys.*, **348**, 642.

- Kaplan, G.H. (ed.) (1981) The IAU resolutions on astronomical reference system, time scales, and the fundamental reference frames, in USNO Circular 163, U.S. Naval Observatory, Washington, D.C.
- Klioner, S.A. (2008) Relativistic scaling of astronomical quantities and the system of astronomical units. *Astron. Astrophys.*, **478**, 951–8.
- McCarthy, D.D. and Petit, G.P. (eds) (2004) *IERS Conventions (2003)*, IERS Technical Note 32, Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.
- Moyer, T.D. (1981) Transformations from proper time on earth to coordinate time in solar system barycentric space-time frame of reference. *Celest. Mech.*, **23**, 33–56, 57–68.
- Nelson, R.A., McCarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L. and Bartholomew, T.R. (2001) The leap second: its history and possible future. *Metrologia*, **38**, 509–29.
- Seidelmann, P.K. and Fukushima, T. (1992) Why new time scales? *Astron. Astrophys.*, **265**, 833–8.
- Standish, E.M. (1998) Time Scales in the JPL and CfA ephemerides. *Astron. Astrophys.*, **336**, 381–4.
- The Astronomical Almanac*, U.S. Government Printing Office, Washington DC.
- Trans. Int. Astron. Union*, Vol. XVI B, *Proceeding of 16th General Assembly Grenoble, 1976* (eds E. Müller and A. Jappel), Association of University for Research in Astronomy, 1977.
- Trans. Int. Astron. Union*, Vol. XVII B, *Proceeding of 17th General Assembly Montreal, 1979* (ed. P. Wayman), Association of University for Research in Astronomy, 1980.
- Trans. Int. Astron. Union*, Vol. XXI B, *Proceeding of 21st General Assembly Buenos Aires, 1991* (ed. J. Bergeron), Kluwer Academic Publishers, 1992.
- Trans. Int. Astron. Union*, Vol. XXIV B, *Proceeding of 24th General Assembly, Manchester, 2000* (ed. H. Rickman), Astron. Soc. of the Pacific, 2001.
- Trans. Int. Astron. Union*, Vol. XXVI B, *Proceeding of 26th General Assembly, Prague, 2006* (ed. K.A. van der Hucht), Cambridge University Press, 2007.

9

Clock Developments

9.1

Introduction

In the English language the word 'clock' is related to the medieval Latin word '*clocca*' and to the French '*cloche*.' Both words mean 'bell,' and refer to the original purpose of clockwork, that is to ring bells at desired time intervals (Whitrow, 1988). Today we use the word 'clock' to refer to any device used to measure the passage of time. These devices may range from the 'clocks' of antiquity that provided measures of time by the flow of water to the most sophisticated clocks of recent times. Generally speaking we can use the word to mean any man-made instrument used to measure time that does not depend on our ability to see the skies. That would then exclude those devices that make use of the length or direction of the Sun's shadow or the apparent motion of the stars in the sky to tell the time.

In the European middle ages, the word 'horologium' was used to refer to any device used to keep time. The continued use of this word throughout the middle ages makes it difficult now to determine who was responsible for the first mechanical clock and when and where it came into existence, because writers of the early middle ages did not distinguish between the different timekeeping devices. Clock development reflects the overall growth in technological ability as well as society's requirements for accurate time to meet everyday needs.

9.2

Keeping Time in Antiquity

Early societies made use of the heavens to meet their need for timing. Measurement of the passage of time in terms of years and days was fairly straightforward. The direction of the Sun's shadow, or its length, are obvious ways to mark time's passage during the day. Instrumentation can be as simple as a stick stuck into the ground or as complex as the most sophisticated sundials. During the night the motion of the stars can be used to indicate both the passage of the seasons as well as the passage of time during the night.

Egyptians were the first to divide the day into shorter time intervals. Their system was to divide the daylight into ten equal parts, to allow an hour each for morning and evening twilight, and to divide the night into twelve equal parts. This provided for a day of 24 hours, but the length of the hours would vary depending on the seasons. In the summer, when the period of daylight would be longer, the daylight hours would be longer and the night time hours shorter. The opposite would occur in winter. Night time would end with the heliacal rising of a particular star or constellation. Heliacal rising refers to the rising of a celestial object at the same time as the Sun rises. The apparent motion of the Sun in the sky, due to the Earth's orbital motion, means that heliacal risings of these asterisms will change throughout the year. Egyptian priests chose to designate the heliacal rising of a different object at ten-day intervals to serve as the markers for the passage of time during the year. These asterisms are known as 'decans' referring to their change at ten-day intervals.

The Egyptian calendar called for 365 days, and so there would be 36 decans with the last 5 days at the end of the year left over for festivals. The identity of all of the decans is not certain, but the star Sirius is known to be one of them. At the time of the heliacal rising of Sirius, or the beginning of the Egyptian year, there are twelve decans which pass during the night. As a result, the nighttime was divided into twelve hours, but with hours of unequal length called 'seasonal' hours. This process resulted in our 24-hour day.

Hellenistic astronomers used a system of hours of equal length for scientific purposes, but the system of seasonal hours continued to be used in everyday practice well into the middle ages. The length of the equal hours was equivalent to the seasonal hours at the time of the equinoxes. These astronomers, following the Babylonian sexagesimal system of counting, further divided the hour into sixty 'firsts' each containing sixty 'seconds,' resulting in our common division of the days into hours, minutes, and seconds (Whitrow, 1988). With the growth of civilizations, the need for time without reference to the sky became a concern, and the difference between the system of seasonal and equal hours posed a problem for the development of clocks through the ages.

9.2.1

Clepsydrae and Water 'Clocks'

The measured flow of water, either into or out of a vessel, was an early means of dealing with the need to have a nonastronomical source of time. A clepsydra (pl: clepsydrae) depends on the uniform flow of water to measure time. The word 'clepsydra' is derived from the Greek words 'to steal' and 'water' and literally means a water thief. It refers to the fact that early forms of clepsydrae were developed from simple siphon devices used to lift water liquids from one vessel to another. Clepsydrae were mentioned in Indian and Chinese texts of the first millennium BC as well as in Babylonian texts. An example of an Egyptian version can be found at Karnak dating to about 1400 BC. This device was even able to allow

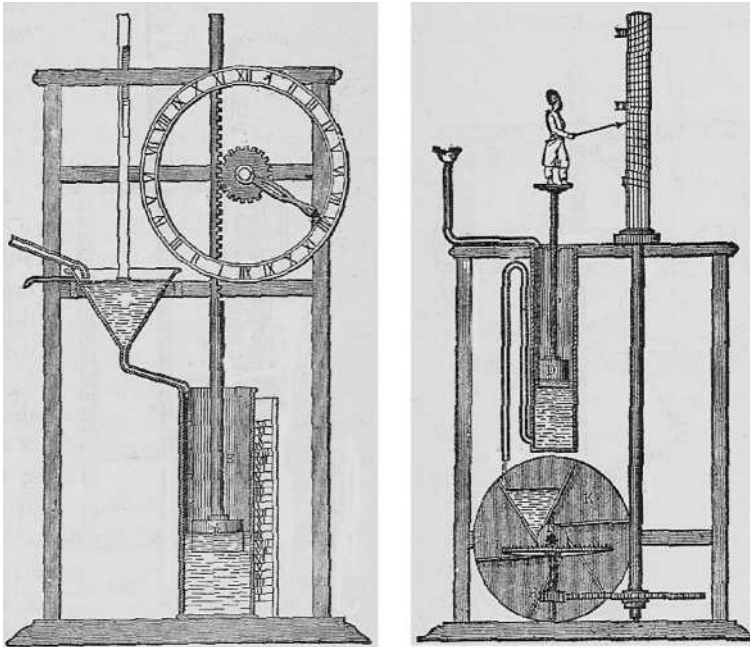


Figure 9.1 Examples of outflow and inflow clepsydrae from Milham (1923).

for the changing length of the hours of day and night during the year. In China clepsydrae were available from about 1500 BC, apparently making use of technology from Babylon.

Greek versions appeared later, when simple outflow timers were used in Athenian law courts to limit the length of speeches. In the third century BC the Greeks developed an inflow clepsydra with a means to regulate the water pressure and thus improve its precision. Such a device was used in the Tower of the Winds in Athens to provide public time (Dohrn-van Rossum, 1996). Figure 9.1 shows two different types. Clepsydrae adapted for use as water ‘clocks’ were used well into the middle ages. Although they were used widely in monasteries, they also found civil applications, and were even used in medical applications of the time.

9.2.2

Other Timekeeping Devices

The flow of water was not the only means of early timekeeping. Burning calibrated candles, or incense sticks, provided one means of keeping time without access to the skies. Candles were popular in the western and Islamic worlds, while incense was used in Asia (Dohrn-van Rossum, 1996).

9.3

The First Mechanical Clocks

The first weight-driven mechanical clocks began to appear in Europe in the latter part of the thirteenth century. There is no documented evidence of the first such clock or the person who built it. Water clocks, based on gear trains regulated by the flow of water, had been in use for some time, and it seems possible that a geared clock based on the regulated fall of a weight might have been a natural progression.

The mechanical clock came into existence with the invention of the verge escapement and foliot regulator (see Figure 9.2). An escapement is a device that controls the continuous motion of the clock's driving mechanism using the periodic motion of its regulator. A weight attached to a pulley would be expected to fall without stopping, but an escapement was introduced to limit the fall to small increments. An arm extending from the 'verge' would then stop the fall and the wheel would apply a rotational impulse to the verge. That releases the wheel until it contacts the second arm of the verge. A crossbar on the verge, called a 'foliot' controls the time it takes for the verge to rotate. The timing is changed by adjusting the positions of the regulating weights. The timing of clocks that used this mechanism were typically about 15 minutes per day. The advent of these verge and foliot clocks led to the gradual decline of water clocks and the development of more sophisticated variants for everyday use.

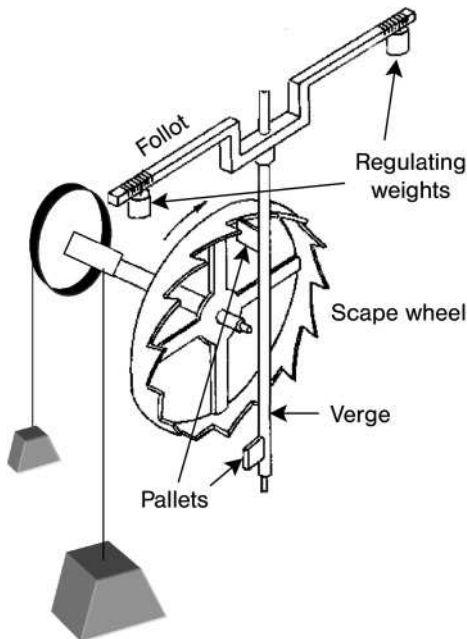


Figure 9.2 Schematic depiction of a verge and Foliot clock.

9.4 Pendulum Clocks

The next important step in clock development occurred with the adaptation of the pendulum to mechanical clock movements. This advance began with the work of Galileo Galilei in the late sixteenth century, and while no longer used as the source of our most precise time, the pendulum clock did serve in that capacity well into the twentieth century.

The period of a pendulum's swing is given by the equation

$$T = 2\pi \sqrt{\frac{l}{g}} \left(1 + \frac{1}{4} \sin^2 \frac{\theta}{2} + \frac{9}{64} \sin^4 \frac{\theta}{2} + \dots \right), \quad (9.1)$$

where l is the length of the pendulum, θ is the maximum semi-amplitude of the pendulum swing, and g is the acceleration due to gravity. The period then depends chiefly on the pendulum's length for small angles and does not depend on its mass. However, the period does depend on the strength of the local gravity vector, and so the period of the motion will depend on the pendulum's location.

9.4.1 Galileo

Galileo's secretary and earliest biographer, Vincenzo Viviani, tells us that Galileo's interest in pendulums began when he was a student in Pisa, where he noticed the motion of a suspended lamp in the cathedral. His notes on pendulums begin in 1588, and in 1602 he wrote to a colleague about his observations, that the swings of a pendulum were isochronous and that pendulums of the same string-length had the same period of swing. In practice, the pendulum was not used as a clock at that time, but it was used to measure heart rates and in scientific experiments. In using the pendulum, its oscillations would be made to continue by an assistant occasionally applying impulses. It was not until 1641, a year before his death, that Galileo began to consider using a pendulum to regulate a clock (see Figure 9.3). Viviani, writing in 1759, 17 years after Galileo's death, described the scene:

One day in 1641, while I was living with him at his villa in Arcetri, I remember that the idea occurred to him that the pendulum could be adapted to clocks with weights or springs, serving in place of the usual *tempo*, he hoping that the very even and natural motions of the pendulum would correct all the defects in the art of clocks. But because his being deprived of sight prevented his making drawings and models to the desired effect, and his son Vincenzio coming one day from Florence to Arcetri, Galileo told him his idea and several discussions followed (Drake, 1978).

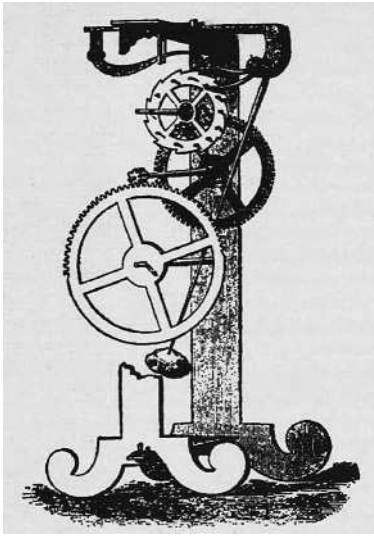


Figure 9.3 Viviani's drawing of Galileo's concept for a pendulum clock from Whitrow (1988).

9.4.2

Huygens

Christiaan Huygens (1629–1695) receives the credit for the next significant advance in clock development with his design for the first successful operational pendulum clock in 1656. Huygens arranged for the patent for his clock to be assigned to a clockmaker named Salomon Coster, who was able to manufacture a few clocks before his death in 1659. His initial effort had an error of less than one minute per day, but subsequent improvements reduced the error to about 10 seconds per day.

Johannes Hevelius, an astronomer who was in the process of building an observatory in Gdansk, had employed a clockmaker, Wolfgang Günther to work on producing a pendulum clock, but he was unable to develop his clock before Huygens' success. In 1673 Hevelius wrote in his *Machinae Coelestis, Pars Prior* (van Leeuwen, 2007 transl.) (Hevelius, 1673):

Around this time, whilst the two pendulum clocks were being worked upon by the clockmaker but which were not completely finished (the clockmaker had little time available due to his work on the larger astronomic instruments), the very distinguished and very scholarly Christiaan Huygens invented similar clocks in 1657. This was also a very successful enterprise, and, a short time later, in 1658, he published an illustration of the pendulum clock, to the great advantage of (scientific) literature and for which I congratulate him. For, this prestigious invention offers an excellent remedy for all the ills of clocks built as yet, as well as solving the problems of inaccuracies which have crept into the escapements as well as the axles, pins, and cog-wheels.

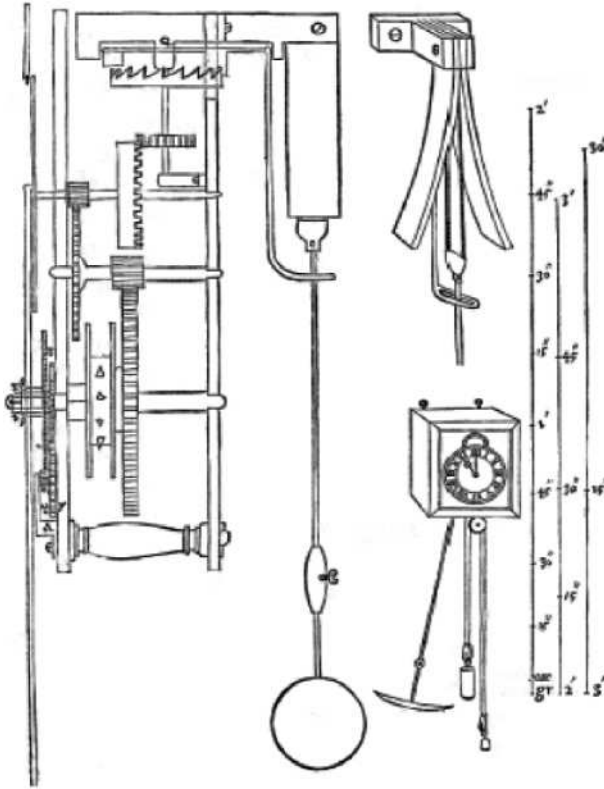


Figure 9.4 Diagram of Huygens' pendulum clock from his *Horologium Oscillatorium*.

The 1658 illustration to which Hevelius refers is contained in a booklet entitled *Horologium* (Huygens, 1658). Huygens pursued his clock development and went on to discover that the circular arc of the pendulum was not truly isochronous. However, if the pendulum bob were constrained to follow the path of a cycloid, it would indeed provide an isochronous regulator. The word 'cycloid' refers to the path of a point on the rim of a rolling wheel. In practice, Huygens did this by employing metal cheeks that constrained the pendulum to follow a cycloidal curve. This development was described in his 1673 publication *Horologium Oscillatorium* (Huygens, 1673). Figure 9.4 reproduces the diagram for Huygens' clock as presented in that publication.

Huygens was also responsible for the first attempt to use time as a means to define a standard of length, when he suggested that the length of a pendulum that produced a particular time could be used as a standard length. However, since the period depends on local gravity, it could not be considered seriously in the definition of a conventional standard.

Huygen's clock still made use of a verge escapement that allowed the weight to drop about five centimeters in one hour. Until then, clocks generally used only an hour hand to indicate time. The improved accuracy of his clock allowed him to

add minute and second hands. However, second hands did not become common until further improvements were made in the escapement.

9.4.3

Pendulum Clock Developments

The first improvement in clock escapements came in the late 1660s, when the first 'anchor' escapement appeared. Its invention is usually credited to William Clement (Whitrow, 1988), but Robert Hooke and Joseph Knibb are also mentioned as possible inventors. Although Huygens' design corrected for the fact that the circular swing of the pendulum was not isochronous, the invention of the anchor escapement allowed pendulum clocks to operate with a reduced swing, making the addition of the metal cheeks less important. The swing of the pendulum could be reduced from around 100° to about 5° . The reduced angle through which the pendulum must swing also resulted in the familiar shape of the longcase or 'Grandfather' clocks.

In the anchor escapement, the pallets look like a ship's anchor, as seen in Figure 9.5. The swinging of the pendulum moves each pallet alternately between the escape wheel teeth. This wheel moves a little in the interval between the time one pallet moves away and the other engages another tooth.

The development of the pendulum did not stop there, however. A well-known family of clockmakers in England, the Fromanteels, sent Johannes Fromanteel to the workshop of Salomon Coster, Huygens' clockmaker, in 1657 to learn the details of the pendulum clock. On his return in 1658, the Fromanteels were able

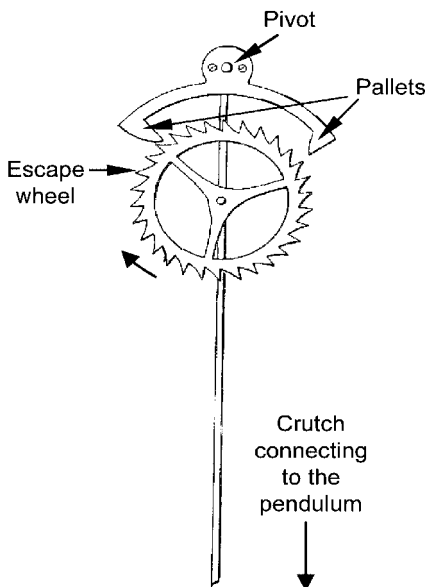


Figure 9.5 Anchor escapement after Whitrow (1988).

to build the first English pendulum clocks. They went on to introduce new features and to set the basic design for longcase clocks into the eighteenth century (Landes, 1983).

The development of pendulum clocks was so successful that by the end of the seventeenth century they could be used for astronomical observations. Thomas Tompion (1639–1713) had finished making two 13-foot pendulums for the new Royal Observatory at Greenwich England in 1676. In 1715 George Graham, Tompion's partner, invented the 'dead-beat' escapement which provided a short impulse to the pendulum when it was nearly vertical. Other improvements over the following centuries dealt with compensating for the effects of temperature on the length of the pendulum, as well as changes in air pressure on the pendulum motion (Landes, 1983).

The most precise clocks were installed in observatories and were referred to as 'astronomical regulators.' These became the devices used to distribute time locally and eventually nationally. Astronomical observations would be made at the observatories in order to maintain the clock's long-term agreement with time determined from the Earth's rotation.

In 1889 the Riefler escapement was invented by Siegmund Riefler. Designed for use with regulators, it made use of a suspension spring to keep the pendulum swinging. This spring not only suspended the pendulum, but it was also able to provide impulses to the pendulum, made of invar and minimally entangled with the escape wheel (Milham, 1923; Sullivan, 2001). Riefler clocks (Figure 9.6) served as the basis for U.S. time beginning in 1909.

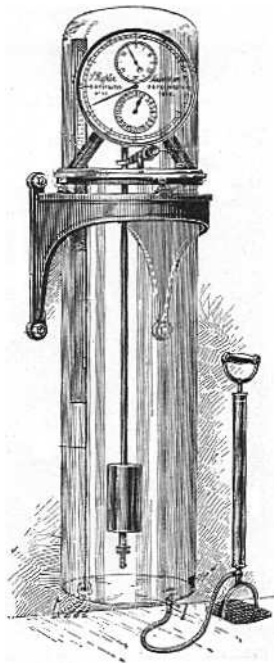


Figure 9.6 Riefler Clock in airtight enclosure from Milham (1923).

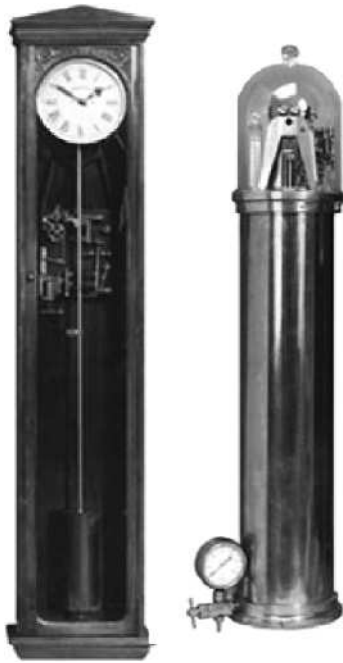


Figure 9.7 Shortt Clock system. The master pendulum is on the right and the slave is on the left (<http://www.electric-clocks.nl/clocks/>).

In 1921, William H. Shortt invented a new pendulum clock that would soon replace the Riefler clocks for use as regulators. It actually made use of two pendulums (Figure 9.7), the master pendulum that actually kept the time and a slave pendulum that provided impulses to the master every 30 seconds (Hope-Jones, 1931). The pulses were delivered electrically to the master pendulum that was again made of invar, to reduce thermal effects, and housed in a partial vacuum. This form of clock, manufactured by the Synchronome Company, became the basis for the U.S. time in 1929 (Sullivan, 2001), and, with an accuracy of 1 milli-second per day, was the first clock capable of measuring variations in the Earth's rotation.

9.4.4

Chronometers

In addition to the development of the regulator clocks, significant progress was also being made in the development of chronometers. The word 'chronometer' is usually applied to high-precision, portable timekeeping devices. They have been used historically to determine longitude at sea or for maintaining railroad schedules. The most significant development in this area of timekeeping was the work of John Harrison (1693–1776). The Longitude Act of Great Britain, passed in 1714, offered a first prize of £20 000 to anyone who could find longitude at sea with accuracy better than half a degree. Harrison proposed to do this by using a clock,

but it meant that he would need to develop a portable clock that was accurate to better than 3 seconds in one day (Harrison, 1767). He succeeded with his chronometer 'H4' and eventually won the prize (Sobel, 1995). In France, Ferdinand Berthoud (1727–1807) also pioneered in producing marine chronometers (Berthoud, 1773).

9.5

Quartz Crystal Clocks

The next significant advance in clock development occurred with the replacement of the Shortt clock by quartz crystal clocks. Timekeeping using these clocks is based on the oscillation of an electrical resonant circuit with current flowing through an inductor with inductance, L , and a condenser with capacitance C (Figure 9.8). The period T of the oscillation is given in hertz by the expression

$$T = 2\pi\sqrt{LC}, \quad (9.2)$$

where L is in henries and C is in farads.

The quartz crystal clock is the result of a long series of steps that began in 1857, when Jules Lissajous was able to sustain the mechanical motion of a tuning fork indefinitely using an electromagnet. A number of developments in electric oscillators occurred over the following years as outlined by Marrison (1948). Time and frequency standards using a tuning fork were gradually developed (Horton and Marrison, 1928; Dye and Essen, 1934), but during that time quartz crystals were also being investigated for possible application in oscillating circuits. The piezoelectric activity and the mechanical and chemical stability of quartz, together with the fact that a relatively small amount of energy was needed to maintain oscillations, made it particularly attractive.

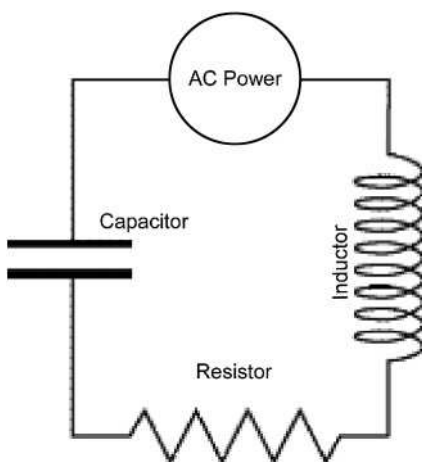


Figure 9.8 Electrical resonant circuit.

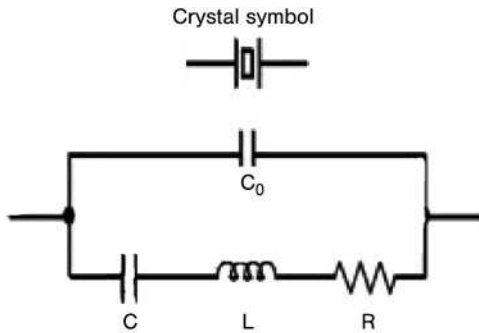


Figure 9.9 Electrical representation of a quartz crystal. The capacitance, C , the inductance, L , and the resistance, R , are inherent in the properties of the crystal itself, and are not separate components within the crystal. The capacitance C_0 is due to the wire connections across the quartz and can be measured physically.

Piezoelectricity is the phenomenon describing the capability of some materials to generate electric potential when mechanically stressed. This ability is called the direct piezoelectric effect. The converse piezoelectric effect refers to the ability of a material to stress or strain when an electric potential is applied. The magnitude of the mechanical motion involved is quite small and only detected optically with high magnification. Some natural, as well as man-made materials, exhibit this capability, including quartz. Jacques and Pierre Curie first demonstrated the direct piezoelectric effect in 1880 (Curie and Curie, 1880, 1882), and through the following years piezoelectricity has found many practical applications.

The natural piezoelectric quality of a quartz crystal means that it will resonate, or ‘ring’, with a frequency depending on its dimensions, producing a voltage across points on its surface to which wires can be attached. Also, it means that the crystal will vibrate the most if the voltage applied across the appropriate spots on its surface is applied with the proper resonant frequency. Quartz crystals can be produced, then, with the correct dimensions to achieve the desired frequencies in the circuits in which they are placed.

In terms of electrical circuits the crystal can be drawn as in Figure 9.9.

Depending on its size and shape, each quartz crystal has its own natural frequency. If it is placed in an oscillating electric circuit that has nearly the same frequency, the crystal will vibrate at its natural frequency, and the frequency of the circuit will become the same as that of the crystal.

In 1917 Alexander McLean Nicolson first used a piezoelectric crystal to control the frequency of an oscillator (Marrison, 1948). Marrison also reports that

The first published quartz-controlled oscillator circuit is reproduced in figure 8A from Cady’s (1922) article. In this oscillator the ‘direct’ and ‘inverse’ piezoelectric effects were

employed separately, making use of two separate pairs of electrodes. The output of a three-stage amplifier was used to drive a rod-shaped crystal at its natural frequency through one pair of electrodes making use of the ‘inverse’ effect, while the input to the amplifier was provided through the ‘direct’ effect from the other pair. The feedback to sustain oscillations in the electrical circuit could be obtained only through the vibration of the quartz rod and hence was precisely controlled by it. Cady’s results were received with widespread interest and were duplicated and continued in many laboratories, which soon resulted in many new discoveries and inventions.

The first quartz clock (Figure 9.10) was built in 1927 by Warren Marrison and J.W. Horton (Horton and Marrison, 1928; Marrison, 1948). Afterwards in the late 1930s, quartz crystal clocks began to replace the pendulum clocks as the time-keeping standards.

If the environment of the quartz oscillator is carefully controlled, its frequency will remain fairly stable. Carefully controlled laboratory quartz crystal clocks may accumulate errors of only a few thousandths of a second in a year.

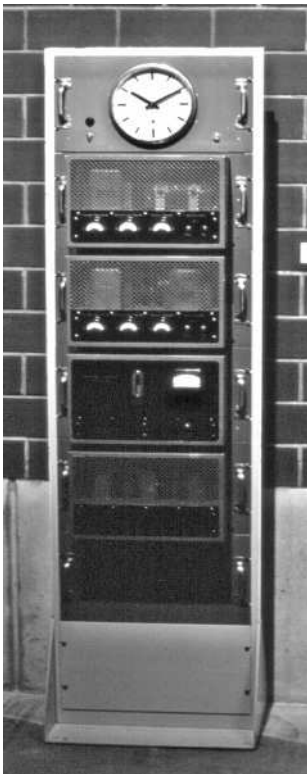


Figure 9.10 The first quartz clock on display at the International Watchmaking Museum, in La Chaux-de-Fonds, Switzerland.

9.6

Clock Performance

Clocks have been, and continue to be, used to provide the time of day in accordance with conventional standards. They also are used to provide time intervals or the time between two events in conventional units. Time interval measurements need not depend on the time of day. To measure either time of day or time interval, clocks depend on the frequency of a repeating phenomenon. In many modern applications the frequency, or time interval, provided by a standard may be more important than the time. For a repeating phenomenon with a period T measured in seconds, the frequency is given as $f = 1/T$ in units of hertz. Frequency comparisons often make use of a dimensionless fractional frequency given by $\Delta f/f$, where Δf is the measured frequency difference between two standards measured at frequency f . We refer to the process of setting two standards to read the same time as ‘synchronization’, and we refer to the process of setting two standards to the same frequency as ‘syntonization.’

The performance of a clock can be characterized by using a number of mathematical terms to describe how well it is, or remains, synchronized or syntonized. For some applications it may not be necessary to relate a standard’s time to a particular reference standard; the stability of its frequency may be the only characteristic of interest. In other applications it may be important to characterize the clock’s ability to reproduce the time provided by a conventionally accepted standard.

9.6.1

Quality (Q) Factor

A term used to describe oscillating systems is the ‘ Q ’ or quality factor. It is defined as the ratio of the total energy in a system to the energy lost per cycle. In general, it refers to the comparison of the time constant describing the decay of an oscillation’s amplitude to the period of the oscillation. This is equivalent to a comparison of the rate at which an oscillation dissipates energy to its oscillation frequency. A high value of Q would mean that an oscillation would die out slowly. A pendulum with a high Q , for example, would be expected to oscillate for a long time, but one with a low value would die out relatively quickly.

For applications in modern timekeeping, the Q factor is related to the width of the resonance phenomenon used to regulate the frequency of the timekeeping device.

$$Q = \frac{f_N}{\Delta f}, \quad (9.3)$$

where f_N is the natural or resonant frequency and Δf is the bandwidth or the range in frequencies that produces oscillation energy at least half of the peak value (see Figure 9.11). A resonant system responds more strongly to a driving frequency closer to its natural frequency. The response of a high- Q system decays much more rapidly as the driving frequency moves away from its natural frequency,

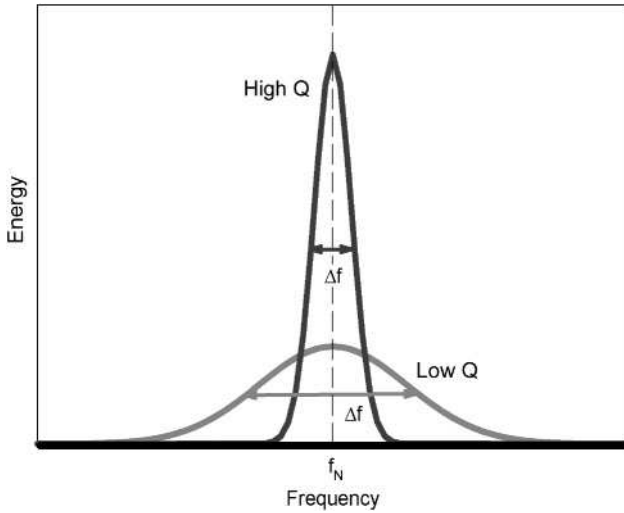


Figure 9.11 Response of a high- Q system (dark) compared to a low- Q system (gray).

so the system with the highest Q would be the most desirable as a frequency standard.

9.6.2

Precision

Precision refers to the ability of repeated measurements to provide the same result. For timekeeping applications we can say that a device is precise if it provides very repeatable measures of a chosen time interval, for example, the length of a second. This can be measured statistically by the familiar standard deviation, σ , of a sample of N measures of the quantity of interest x_i , with mean.

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}}. \quad (9.4)$$

9.6.3

Accuracy

In contrast to precision, accuracy describes the ability of measurements to conform to a standard value. Measurements could be accurate, but not precise, and vice versa (see Figure 9.12). For time-keeping applications we might call a device accurate if it reproduces the conventionally accepted standard value of a time interval after repeated measurements.

Statistically this can be measured by the root mean square (RMS) of measures (x_i) of the difference in frequency, or time, between the clock in question and a

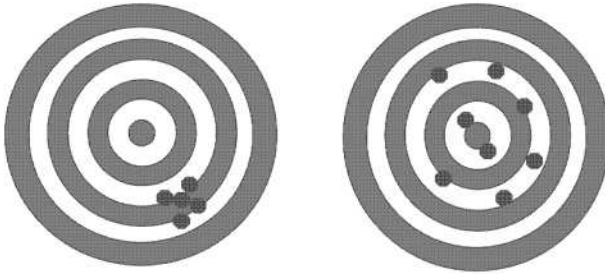


Figure 9.12 Examples of precision without accuracy (left) and accuracy without precision (right).

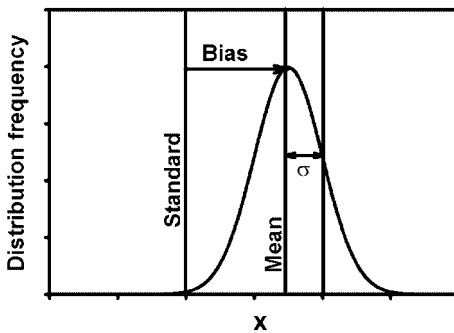


Figure 9.13 Representative distribution plot showing the relationship between the mean, bias, and standard deviation of measurements of the difference between a clock reading and the conventional standard.

conventionally adopted standard, or by determining the bias and standard deviation of those measures (see Figure 9.13).

$$RMS = \sqrt{\frac{\sum_{i=1}^N (x_i)^2}{N}}. \quad (9.5)$$

Figure 9.14 shows the improvement in timekeeping accuracy with time.

9.6.4

Stability

Stability refers to the ability of a standard to maintain its synchronization or syn-
 tonization over time. A stable clock would be one that would produce the same
 measures over a range of time intervals. It does not necessarily have to be
 accurate.

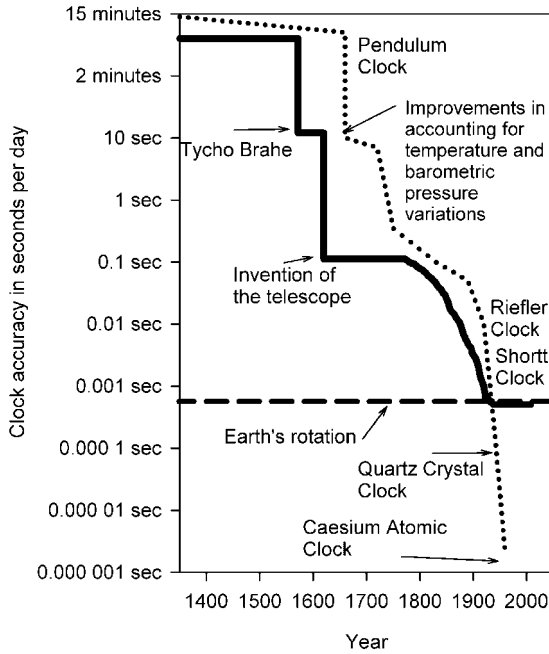


Figure 9.14 The improvement in timekeeping accuracy with time. The solid line shows the accuracy of astronomical optical measurements. The dotted line shows the accuracy of clock measurements. The dashed line shows the accuracy of the Earth rotation to provide an accurate time scale.

If we model the time produced by a clock to be

$$t_c = t + x(t), \quad (9.6)$$

where t is the accurate 'true' time and $x(t)$ is the time varying error in time (or phase of an output sinusoidal signal), then the error in the clock frequency is given by

$$y(t) = \frac{dx(t)}{dt}, \quad (9.7)$$

Statistically we can represent the variance of the errors in time and frequency, respectively, in terms of a Fourier representation by

$$\text{var}(x) = \int_1^\infty S_x(f) df \quad \text{and} \quad \text{var}(y) = \int_1^\infty S_y(f) df, \quad (9.8)$$

where f is the Fourier frequency and $S_x(f)$ and $S_y(f)$ are the spectral densities of the errors in time and frequency, respectively. The spectral densities of time and frequency are related by the expression

$$S_y(f) = 4\pi^2 f^2 S_x(f). \quad (9.9)$$

Clock stability is modeled empirically considering five different types of noise: white phase noise, flicker phase noise, white frequency noise, flicker frequency noise, and random walk frequency noise. The power spectrum of clock noise can be considered, then, to be the sum of these components. For example,

$$S_y(f) = \sum_{\alpha=-2}^{\alpha=+2} h_{\alpha} f^{\alpha}, \quad (9.10)$$

where the individual contributions in the frequency domain are:

$$\begin{aligned} \text{white phase noise : } & h_2 f^2 \\ \text{flicker phase noise : } & h_1 f^1 \\ \text{white frequency noise : } & h_0 f^0 \\ \text{flicker frequency noise : } & h_{-1} f^{-1} \\ \text{random walk frequency noise : } & h_{-2} f^{-2} \end{aligned} \quad (9.11)$$

In the time domain we can refer to Figure 9.15 for a schematic representation of the effects of these types of noise on the observed clock errors as a function of time. Figure 9.16 shows the spectral appearance of the different noise types.

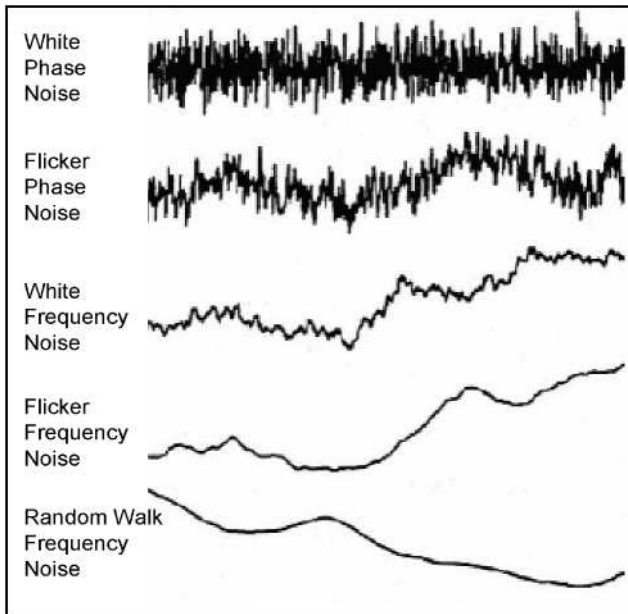


Figure 9.15 Appearance of different noise types in time residuals.

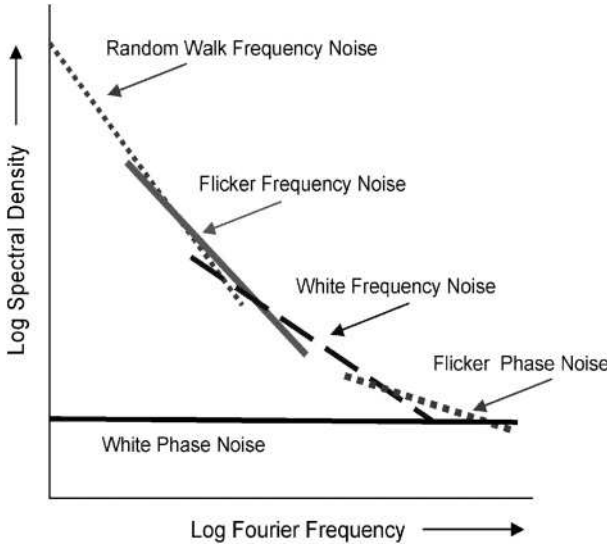


Figure 9.16 Spectral appearance of noise types.

Quantitatively, stability can be measured in practice by the two-sample or Allan variance, whose square root may be called the two-sample or Allan deviation,

$$\sigma_y(\tau) = \sqrt{\frac{\sum_{i=1}^{M-1} (y_{i+1} - y_i)^2}{2(M-1)}} \quad \text{or} \quad \sigma_y(\tau) = \sqrt{\frac{\sum_{i=1}^{N-2} (x_{i+2} - 2x_{i+1} + x_i)^2}{2(N-2)\tau^2}}, \quad (9.12)$$

where y_i is a set of M frequency offset measurements, or x_i is a set of N time difference measurements, all the data being equally spaced at intervals of τ seconds. Instead of subtracting the mean from each data point in the summation, as is the case with the standard deviation, the Allan variance subtracts the following data point. This eliminates the possible contribution of any systematic offset in frequency (Levine, 1999; Lombardi, 1999).

In terms of the five noise types, if we plot $\sigma_y^2(\tau)$ as a function of τ , then we find the following proportional relationships:

$$\begin{aligned} \text{white phase noise:} & \quad H_2 \tau^{-2} \\ \text{flicker phase noise:} & \quad H_1 \tau^{-2} \\ \text{white frequency noise:} & \quad H_0 \tau^{-1} \\ \text{flicker frequency noise:} & \quad H_{-1} \tau^0 \\ \text{random walk frequency noise:} & \quad H_{-2} \tau^1 \end{aligned} \quad (9.13)$$

A plot of $\sigma_y^2(\tau)$ versus the sampling interval would typically appear as in Figures 9.17 and 9.18 shows how quartz crystal clocks would typically appear in a plot of Allan Deviation.

Figure 9.17 shows that the Allan Variance is unable to distinguish between white phase noise and flicker phase noise because both appear with a τ^{-2} dependency. Another concern in using the Allan Variance in describing clock stability is the effect of a deterministic clock variation. The Allan Variance assumes the clock data to be without any coherent variation. The presence of a coherent signal in the data will affect the description of the noise process. To remedy the first problem Allan devised the Modified Allan Variance given by Eq. (9.14) for a set of phase data x_i sampled over N equal intervals of length τ_0 (Allan and Barnes, 1981).

$$\text{Mod}\sigma_y^2(\tau) = \frac{1}{2n^4\tau_0^2(N-3n+1))} \sum_{j=1}^{N-3n+1} \left[\sum_{i=j}^{n+j-1} (x_{i+2n} - 2x_{i+n} + x_i) \right]^2. \quad (9.14)$$

This statistic is able to distinguish between white phase noise and flicker phase noise as shown in Figure 9.18. Essentially it replaces the individual sample in the expressions in Eq. (9.12) with averages over adjacent intervals.

A related statistic called ‘Time Variation’, or TVAR, is sometimes used to describe clock noise. It is given by the expression

$$\text{TVAR} = \sigma_x^2 = \frac{\tau^2}{3} \text{mod}\sigma_y^2. \quad (9.15)$$

Other similar statistics have also been developed to characterize the time varying behavior of clocks and frequency standards. These include Total Variance or Totvar (Howe and Greenhall, 1997; Greenhall, Howe and Percival, 1999; Howe

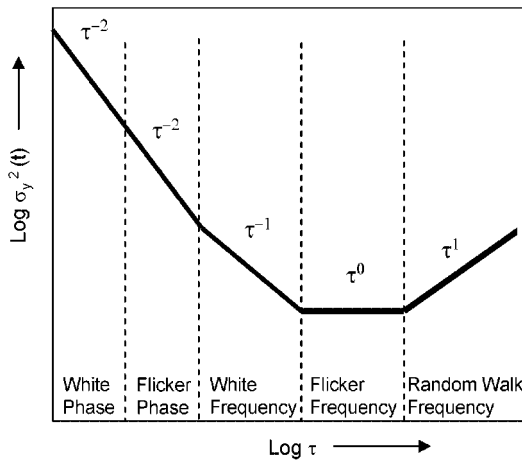


Figure 9.17 Representation of different noise processes in a plot of the Allan Variance as a function of sampling time.

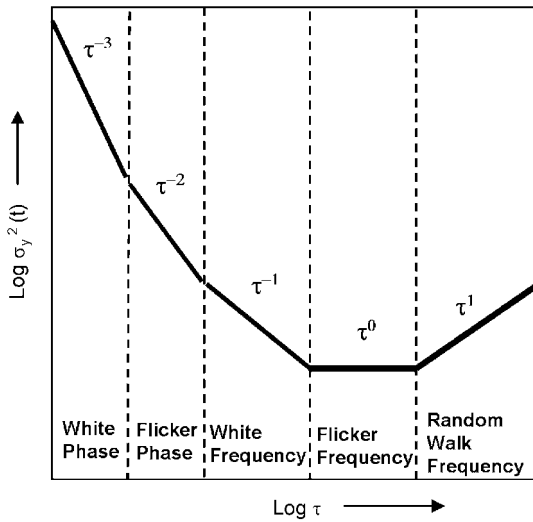


Figure 9.18 Representation of different noise processes in a plot of the Modified Allan Variance as a function of sampling time.

and Vernotte, 1999) and the Hadamard Variance utilizing the second differences in frequency samples, or the third difference in phase samples (Baugh, 1971; Hutsell, 1995). This statistic has the advantage of being insensitive to frequency drift. Associated with the latter are the Modified Hadamard Variance (Bregni and Jmoda, 2006) and the Total Hadamard Variance (Howe *et al.*, 2005).

References

- Allan, D. and Barnes, J.A. (1981) A modified 'allan variance' with increased oscillator characterization ability, Proc. 35th Ann. Freq. Control Symposium, USAERADCOM, Ft. Monmouth, NJ, 1990, in *Characterization of Clocks and Oscillators*, Technical Note 1337 (eds D.B. Sullivan, D.W. Allan, D.A. Howe, F.L. Walls), National Institute of Standards and Technology, Boulder, CO, pp. TN254–7.
- Baugh, R.A. (1971) Frequency modulation analysis with the hadamard variance. *Proceedings of the 25th Annual Symposium on Frequency Control*, pp. 222–5.
- Berthoud, F. (1773) *Traité des horloges marines*, J.B.G. Musier fils, Paris.
- Bregni, S. and Jmoda, L. (2006) Improved estimation of the hurst parameter of long-range dependent traffic using the modified hadamard variance. *Proceedings of the 2006 IEEE International Conference on Communications*, pp. 566–72.
- Cady, W.G. (1922) The piezoelectric resonator. *Proc. Inst. Radio Eng.*, **10**, 83–114.
- Curie, J. and Curie, P. (1880) Developpement par pression, de l'électricité polaire dans les cristaux hémihédres a faces inclinées. *CR*, **91**, 294.
- Curie, J. and Curie, P. (1882) Déformations électrique du quartz. *CR*, **95**, 194–7.
- Dohrn-van Rossum, G. (1996) *History of the Hour*, translated by T. Dunlap, University of Chicago Press, Chicago, IL.

- Drake, S. (1978) *Galileo at Work: His Scientific Biography*, University of Chicago Press, Chicago, IL.
- Dye, D.W. and Essen, L. (1934) The valve maintained tuning fork as a primary standard of frequency. *Proc. R. Soc. London A*, **143**, 285–306.
- Greenhall, C.A., Howe, D.A. and Percival, D.B. (1999) Total variance, an estimator of long-term frequency stability. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, UFFC-46, 1183–91.
- Harrison, J. (1767) *The Principles of Mr. Harrison's Timekeeper*, W. Richardson and S. Clark and sold by J. Nourse, and Mess. Mount and Page, London.
- Hevelius, J. (1673) *Machinae Coelestis*, Pars Prior, Gdansk.
- Hope-Jones, F. (1931) *Electric Clocks*, N.A.G. Press, London.
- Horton, J.W. and Marrison, W.A. (1928) Precision determination of frequency. *Proc. Inst. Radio Eng.*, **16**, 137–54.
- Howe, D.A. and Greenhall, C.A. (1997) Total variance: a progress report on a new frequency stability characterization. *Proc. 29th Ann. PTTI Systems and Applications Meeting*, pp. 39–48.
- Howe, D.A. and Vernotte, F. (1999) Generalization of the total variance approach to the modified Allan variance. *Proceedings of 31st PTTI Systems and Applications Meeting*, pp. 267–76.
- Howe, D.A., Beard, R.L., Greenhall, C.A., Vernotte, F., Riley, W.J. and Pepler, T.K. (2005) Enhancements to GPS operations and clock evaluations using a 'total' hadamard deviation. *IEEE Trans. Ultrason. Ferroelec. Freq. Control*, **52**, 1253–61.
- Hutsell, S.T. (1995) Relating the hadamard variance to MCS kalman filter clock estimation. *Proceeding of 27th PTTI Meeting*, pp. 291–302.
- Huygens, C. (1658) *Horologium*, transl. by E. L. Edwardes, 1970, *Horologium*, by Christiaan Huygens, 1658. *Antiq. Horol.*, **7**, 35–55.
- Huygens, C. (1673) *Horologium oscillatorium sive de motu pendulorum ad horologia aptato demonstrationes geometricae*, Paris, 1673; English translation: *The pendulum clock or geometrical demonstrations concerning the motion of pendula as applied to clocks*, Iowa State University Press, available online at http://historical.library.cornell.edu/kmoddl/toc_huygens1.html and translation available online at <http://www.17centurymaths.com/contents/huygens/horologiumpart5.pdf>.
- Landes, D.S. (1983) *Revolution in Time, Clocks and the Making of the Modern World*, Belknap Press of the Harvard University Press, Cambridge, MA and London.
- van Leeuwen, P. (2007) http://www.antique-horology.org/_Editorial/Hevelius/.
- Levine, J. (1999) Introduction to time and frequency metrology. *Rev. Sci. Instrum.*, **70**, 2567–95.
- Lombardi, M.A. (1999) *Fundamentals of Time and Frequency*, in *The Mechatronics Handbook*, CRC Press, Boca Raton, FL, pp. 17-1–17-18.
- Marrison, W.A. (1948) The evolution of the quartz crystal clock. *Bell Syst. Tech. J.*, **XXVII**, 510–88.
- Milham, W.I. (1923) *Time & Timekeepers Including the History, Construction, Care, and Accuracy of Clocks and Watches*, Macmillan, New York and London.
- Sobel, D. (1995) *Longitude, the True Story of a Lone Genius who Solved the Greatest Scientific Problem of His Time*, Walker and Company, New York.
- Sullivan, D.B. (2001) Time and frequency measurement at NIST: the first 100 years. *2001 IEEE Int'l Frequency Control Symposium*, National Institute of Standards and Technology.
- Whitrow, G.J. (1988) *Time in History*, Oxford University Press, Oxford.

10

Microwave Atomic Clocks

10.1

Beyond Quartz-Crystal Oscillators

Although quartz-crystal oscillators had provided a major advance in timekeeping, it was apparent that there were limitations to that technology. These devices could provide frequency with a precision of about 10^{-10} , but going beyond that would be a challenge. Operationally, fundamental mode crystals could be made to provide frequencies up to 50 MHz. Higher frequencies capable of providing more precise time were possible using overtones but not commonly used. Aging and changes in environment, including temperature, humidity, pressure, and vibration, affected the frequency of the crystal, so systems were designed to attempt to compensate for these problems, including temperature-compensated crystal oscillators (TCXO), oven-controlled crystal oscillators (OCXO), and micro-computer-controlled crystal oscillators (MCXO). Today, systems are being commonly used that utilize signals from navigation satellite systems, such as the Global Positioning System (GPS), to discipline crystal oscillators in order to correct for these effects.

To make a significant advance in precision timekeeping of laboratory standards a fundamental change was required. In his autobiographical account, *Time for Reflection*, Louis Essen, the maker of the first operational cesium (symbol Cs) atomic clock, writes (Essen, 2000):

‘The only possible alternative appeared to be a natural periodicity within the atom. The science of optical spectroscopy had given a picture of the atom as a miniature solar system having a central nucleus with a number of electrons revolving round it in permitted orbits. When an electron jumped from one orbit to another, light of a specific frequency was emitted or absorbed according to whether the jump was to a higher or lower energy state. Well-known examples are the mercury and sodium lamps that illuminated the streets. The frequencies of these optical spectral lines are too high to be measured directly in terms of quartz clocks but are calculated from their wavelength, which

can be measured, and the velocity of light, velocity being the product of wavelength and frequency.

‘The development of microwave techniques during the war provided spectroscopists with a powerful new tool and enabled them to study the response of atoms to electromagnetic waves covering a whole new band of frequencies. Atoms of the alkali metals were of particular interest because they have a single electron in the outermost orbit and, therefore, give the simplest spectrum. The results were brilliantly interpreted and led to the assumption that the outer electron and nucleus were spinning in either the same or opposite direction and that the two conditions represented states having slightly different energies. Transitions between them were accompanied by the emission or absorption of a (quantum of) radiation in the microwave region of the spectrum. The significance of this from our point of view is that the frequency can be measured in terms of the quartz standards, and it becomes potentially possible to define the second as the time occupied by a certain number of cycles of an atomic spectral line. It was suggested by Rabi that a spectral line of an isotope of caesium might be suitable. Its frequency was near 10^{10} Hz in a band used for radar.’

The story of atomic clocks doesn’t begin with Essen, however. It actually goes back to the late nineteenth century, but to trace that history it is important to first understand the physics of using atoms to determine time and frequency.

10.2

Physics of Atomic Clocks

Elements are composed of atoms, each of which has an identical structure that characterizes that particular element. Isolated atoms of an element have identical energy levels, and changes of those energy levels occur when the atomic electrons interact with electromagnetic fields. Energy level transitions are related to the frequency of the relevant electromagnetic field, causing a transition by the expression,

$$\nu = \frac{E_2 - E_1}{h}, \quad (10.1)$$

where E_1 and E_2 are the energies of the levels involved, ν is the frequency of the magnetic field, and h is Planck’s constant. This expression shows that it is possible to provide a frequency from atomic energy level transitions that might be useful for time-keeping purposes.

Atomic energy levels take on discrete values and are classified according to a series of distinct physical states. The principal levels, which are characterized by

the principal quantum number n , describe the radii of electron orbitals about the nucleus. These levels describe the largest atomic energy separations. The elements of principal interest for atomic clocks are those with just a single electron in their outer orbital shell. The alkali metals (including rubidium and cesium) all have one unpaired electron in the outer shell. Their inner shells are either empty or completely full. Similarly, ions with only one electron in the outer shell can be used in atomic timekeeping. The principal quantum numbers for the energy levels used in the most common clock transitions are $n = 1$ for hydrogen, $n = 5$ for rubidium, and $n = 6$ for cesium and singly ionized mercury.

While n characterizes the radii of the electron orbitals about the nucleus, the principal energy levels are subdivided further as a result of the quantization of the angular momentum of the electrons. These levels are characterized by the quantum number l . For the transitions of interest in atomic timekeeping, $l = 0$ (the ground state) and $l = 1$. Electrons are further distinguished by another property called spin, characterized by the quantum number s . This number can only take on the values of $+\frac{1}{2}$ or $-\frac{1}{2}$. Consequently, the total electron angular momentum, which is the sum of the orbital and spin angular momenta of the electron, is characterized by the quantum number J given by $l + \frac{1}{2}$ or $|l - \frac{1}{2}|$. For the ground state, then, $J = \frac{1}{2}$, and in the first excited state ($l = 1$) $J = \frac{1}{2}$ or $J = \frac{3}{2}$. The nucleus of the atom also possesses spin with an associated quantum number I , which, for alkali metals, is given by an odd multiple of $\frac{1}{2}$. For the elements of interest for atomic clocks these numbers are $I = \frac{1}{2}$ for hydrogen and singly ionized mercury, $I = \frac{3}{2}$ for rubidium, and $I = \frac{7}{2}$ for cesium.

The total angular momentum of the atom, then, is the vector sum of the angular momenta of the electron and the nucleus and characterized by the quantum number F . It can only take on integral values $|I - J| \leq F \leq I + J$. In the presence of a magnetic field the Zeeman effect splits each hyperfine level, characterized by the quantum level F , into $2F + 1$ sub-levels characterized by the quantum number m_F , which can take on values $-F \leq m_F \leq F$. The levels characterized by $m_F = 0$ are referred to as ‘clock transitions’ because they have very small variations in the values of their energy as the magnetic field varies (Audoin and Guinot, 2001). Figure 10.1 shows an energy level representation of the fine and hyperfine structure of the outer ($n = 6$) electron of the cesium atom as an illustration. The hyperfine structure is further subdivided into Zeeman sublevels caused by the application of constant magnetic fields.

A common notation, also shown in Figure 10.1 is used to specify these levels. An upper case letter is used to specify the orbital angular momentum number l . So for $l = 0$ the letter S is used and for $l = 1$ P is used. A superscript index preceding the uppercase letter specifies the quantity $2s + 1$, and a subscript index after the letter corresponds to the value of J .

The accuracy of the frequency corresponding to any transition is limited by the Heisenberg uncertainty principle:

$$\Delta E \Delta t \geq \frac{h}{2\pi}. \quad (10.2)$$

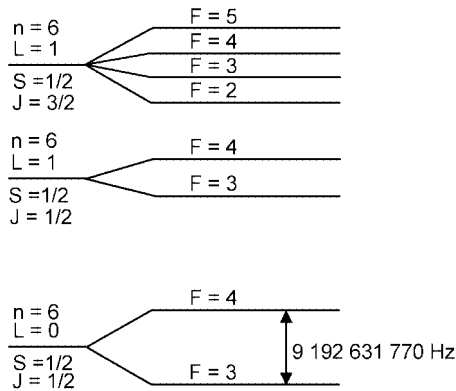


Figure 10.1 Energy level representation of the fine (left) and hyperfine (right) structure of the outer ($n = 6$) electron of the cesium atom.

Since $E = h\nu$, $\Delta\nu \Delta t \geq 1$, implying that long observation times would lead to a smaller frequency uncertainty. In practice, the spectral lines corresponding to the energy level transitions are broadened both for physical reasons and because of the finite resolution of the electronics involved. The most significant physical causes are due to temperature and pressure broadening.

10.3

General Structure of Atomic Clocks

In general, atomic clocks are constructed by combining (i) an oscillator with excellent short-term stability such as a quartz oscillator, (ii) an atomic resonator capable of being locked to the oscillator in an electronic circuit, and (iii) the means to use the frequency of the electronic circuit to provide an indication of time. Figure 10.2 shows the concept schematically.

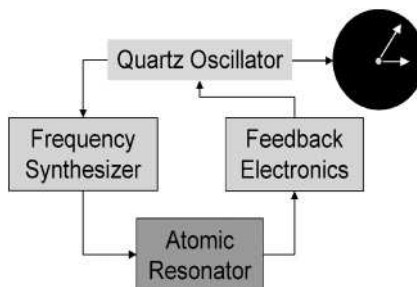


Figure 10.2 Atomic clock concept.

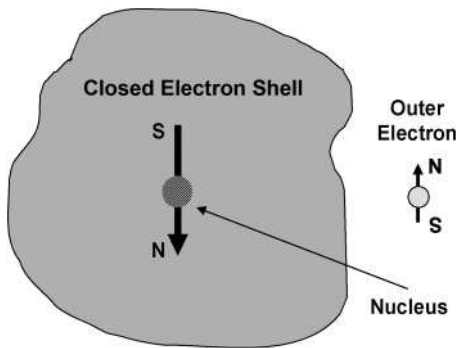


Figure 10.3 Hydrogen-like atom with one outer electron whose spin vector is oriented opposite to the nuclear spin vector.

Atomic frequency standards are commonly based on hyperfine transitions of hydrogen-like atoms, such as rubidium, cesium, and hydrogen. These atoms have a single unpaired electron in a symmetric orbit where there is no orbital angular momentum and no fine structure. Their fundamental state breaks down into only two hyperfine levels because of the interaction with the nucleus' magnetic spin. These provide transition frequencies that can be used conveniently in electronic circuitry, for example, 1.4 GHz for hydrogen, 6.8 GHz for rubidium, and 9.2 GHz for cesium. These hydrogen-like (or alkali) atoms have structures as shown in Figure 10.3. The energy level transitions then refer to a change in the quantum number F as shown in Figure 10.4. They correspond to the spin-spin interaction between the atomic nucleus and the outer electron in the ground state of the atom and are called the ground state hyperfine transitions. As mentioned above, the 'clock transition' is the transition between the least magnetic-field-sensitive sub-levels. To minimize the possibility of other transitions occurring, a constant magnetic field, called the 'C-field,' is applied in the resonator.

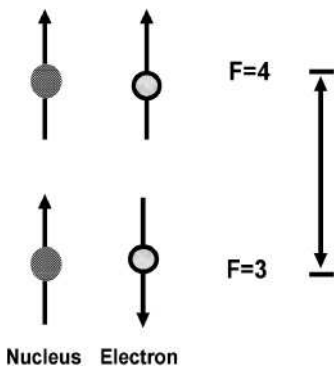


Figure 10.4 Representation of the energy level transition in the hydrogen-like atoms.

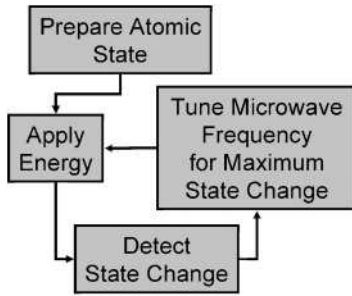


Figure 10.5 Schematic representation of an atomic resonator designed to provide the frequency corresponding to an energy level transition.

An atomic resonator, designed to provide a frequency ν corresponding to that transition, is shown schematically in Figure 10.5. First, a maximum number of atoms in the ground state are prepared and exposed to microwave energy with a frequency close to the appropriate frequency corresponding to the desired transition. This frequency is then varied to achieve the maximum number of atoms in the higher energy level.

To be useful for timekeeping purposes, it is important to minimize the fractional frequency measurement uncertainty $\Delta f / f_0$ where Δf is the error in the frequency measurement and f_0 is the resonance frequency. This quantity is inversely proportional to the interrogation time, or the time during which the radiation is applied before measurements are made, and the square root of the number of atoms. Assuming no limitation to the detection efficiency, the Allan Deviation of a frequency standard can be related to the clock transition frequency f_0 , the interrogation time T , and the number of atoms N by the expression

$$\sigma_y(\tau) = \frac{1}{f_0 \sqrt{NT\tau}}. \quad (10.3)$$

So this fact makes it desirable for atomic clocks to make use of a high resonance frequency with a large number of atoms and a long interrogation time.

A number of physical effects can contribute to instability in atomic clocks. These include changes in temperature, humidity, atmospheric pressure, vibration, presence of outside electric and magnetic fields, noise in the circuit electronics, and change in the gravity field. Great care is exercised in the construction of atomic clocks to minimize these effects along with a series of other smaller effects. For example, extensive magnetic shielding is used to reduce the effect of the Earth's magnetic field.

Black-body radiation due to the temperature of the surrounding environment contributes a systematic shift in the frequency that is measured by an atomic clock. Because the atoms involved are in random motion, the observed frequency is broadened by the Doppler Effect. Restricting the space in which the atoms interact limits this problem. Nevertheless, first-order Doppler remains a source of error. In addition to this effect there is a second-order Doppler Effect caused by relativistic time dilation effects for atoms moving with respect to the apparatus. Only

slowing the speeds of the atoms can help with this problem. Collisions of the atoms with the confining chamber contribute to broadening of the observed frequency, and the signal-to-noise ratio contributes a random noise component. Tuning of the cavity in which the interaction occurs can affect the apparent frequency. This effect, called ‘cavity pulling,’ can be helped by going to more narrow resonance peaks.

10.4

Development of Atomic Clocks

The story of the development of atomic clocks does not begin with an atomic resonator using the hydrogen-like atoms, however. The first atomic clock used the ammonia molecule instead. Work in the area of detecting spectral lines corresponding to energy level transitions had focused on the optical frequencies until 1934, when C. E. Cleaton and N. H. Williams were able to observe a change in state of the ammonia molecule that provided a radio frequency at 24 GHz (McCoubrey, 1996). Research in radar during World War II led to further development of microwave spectroscopy. In 1948 Harold Lyons of the United States National Bureau of Standards began his work in developing this spectral line for use in timekeeping, and in 1949 he was able to report success (Lyons, 1949). This clock, shown in Figure 10.6, used a quartz crystal oscillator to provide a frequency



Figure 10.6 The first ammonia standard. The inventor, Harold Lyons, is on the right and the Director of the National Bureau of Standards, Edward Condon, is on the left (Lombardi, Heavner and Jefferts, 2007).

standard controlled by the vibrations of excited ammonia molecules. It reached a stability of the order of 1×10^{-7} , but was limited by Doppler (thermal) and pressure broadening of the spectral line. A second version that reached 2×10^{-8} was not able to out-perform the best quartz crystals of the time (Lombardi, Heavner and Jefferts, 2007), but these devices did point the way to future atomic clocks.

10.4.1

Cesium

The standard for modern timekeeping, the *Système International* second, is based on the cesium atom. Cesium is a nonradioactive element having atomic number 55 and atomic weight 133. It is part of the alkali metals (group I of the periodic table) and reacts violently with water and oxygen. The development of atomic clocks is based on the work of Issac Rabi and colleagues (1939), who pioneered the method of molecular beam magnetic resonance, eventually leading to the cesium beam frequency standard. The first measurements of the hyperfine transition that has become the standard for timekeeping were reported by S. Millman and P. Kusch (1940). In his 1945 Richtmeyer lecture Rabi first mentioned the possibility of developing atomic clocks (Ramsey, 1993). A very significant development in the story is the invention in 1949 by Norman Ramsey (1949, 1950) of a method to use separated oscillatory fields to excite the atoms in an experiment involving the resonance of molecular hydrogen. Instead of distributing the energy

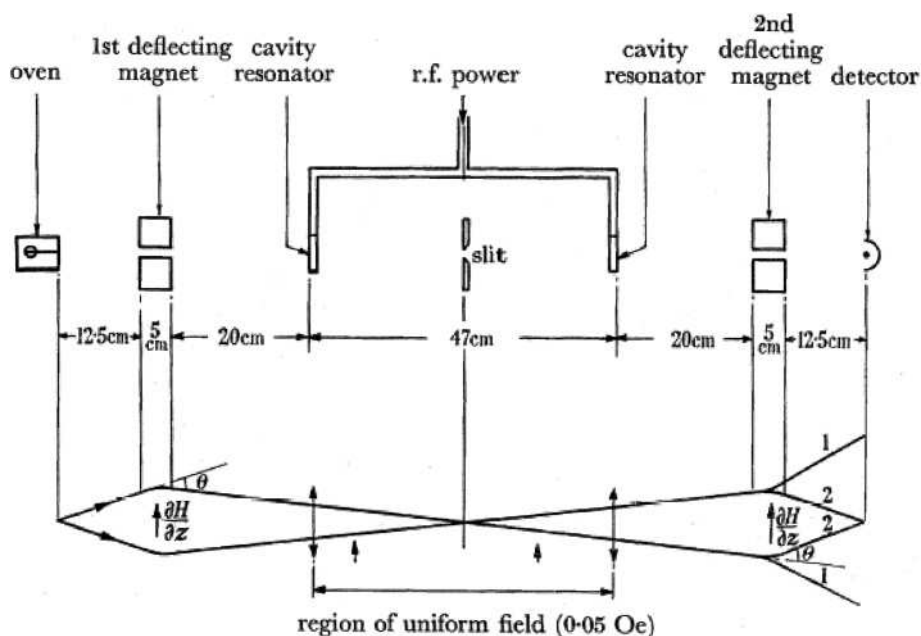


Figure 10.7 Components of the cesium beam chamber by Essen and Parry (1957).

over the entire transition region, Ramsey concentrated it in two coherently driven oscillating fields, one at the beginning and one at the end of the transition region. Using this method provides narrower resonance peaks that are not broadened by inhomogeneities in the field. It also enables the transition region to be larger and permits improved sensitivity through the introduction of relative phase shifts between the two regions (Ramsey, 1983).

In 1952, Sherwood, Lyons, McCracken and Kusch developed the concept of an atomic beam clock and proposed a plan for such a device (Sherwood *et al.*, 1952). J. R. Zacharias in 1954 attempted unsuccessfully to use cesium in a fountain arrangement to provide a frequency source (Ramsey, 1983, 2005). Finally in 1955, L. Essen and J. V. L. Parry at National Physical Laboratory (NPL) in Teddington, UK, produced the first operational cesium beam atomic clock (Essen and Parry, 1957). The components of the beam chamber are shown in Figure 10.7. Figure 10.8 shows a line drawing of the completed standard. Figure 10.9 shows Parry and Essen with the original clock.

In Essen's explanation in his *Time for Reflection* he writes:

'Atoms leave the oven through a narrow slit and pass between the pole pieces of a powerful magnet which is shaped to give a nonuniform field. They follow various paths according to their initial direction, velocity, and energy state. Only two paths are shown. A few of the atoms are selected by the slit half way along the path and continue through the pole pieces of a second magnet, which is the same as the first and increases the deflections in the same direction and away from the centre line. A weak radio field is applied in the space between the magnets and, when its frequency and strength are exactly right, the atoms jump to the other state, those initially in the low energy state absorbing energy from the field and, strangely enough, those in the high energy state being induced to emit energy, so that they are all reversed. Their deflections in the second magnet are also reversed and they are deflected back to the centre line where they strike the detector. This is a hot tungsten wire which imparts a charge to the atoms which boil off as charged particles and are attracted to electrodes and, after enormous amplification, measured as an electric current. Atoms which are not in the two states concerned are not deflected at all and give a steady signal which is useful for lining up the apparatus. The beam strength increases by ten per cent when transition occurs. The components are all contained in a metal pipe about 150 cm long evacuated as completely as possible.'

Cesium was chosen as the element to be used in this effort because:

1. its fundamental state has only two hyperfine levels, just like all the alkali atoms (see Figure 10.1), and at room temperature all cesium atoms are in that fundamental state;

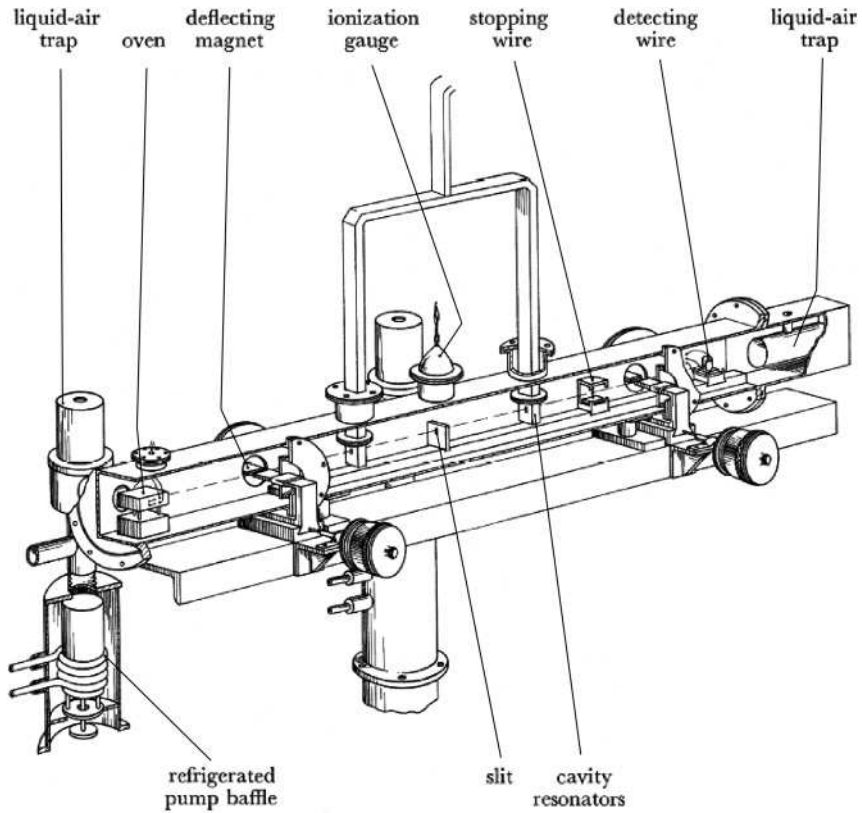


Figure 10.8 Cesium beam chamber by Essen and Parry (1957).

2. in the ($F = 3$ to $F = 4$) transition, atoms in the $F = 4$ state stay there for a long time compared to the observation time;
3. the transition frequency is easily detectable using the electronic equipment in use at that time;
4. it is relatively insensitive to electric fields;
5. it is less expensive than other alkali elements.

10.4.1.1 Calibration of the Cesium Frequency

In order to make practical use of the successful operation of the NPL cesium clock it was necessary to establish a relation between the time derived from the cesium clock and the standard time in use at that time. A preliminary effort was accomplished in June 1955, when Essen and Parry (1955) calibrated the frequency of the transition in terms of the time scale maintained at the Royal Greenwich Observatory, which they could relate to the second of UT2. Concurrently the astronomical community was discussing the use of Ephemeris Time as a uniform time scale to replace UT2 as the time scale of choice for those needing a time scale independent of the Earth's variable rotational speed.

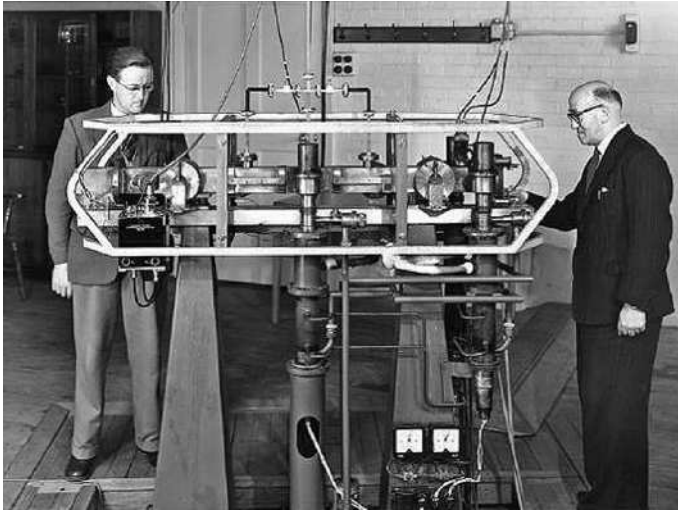


Figure 10.9 Parry (left) and Essen (right) with the first operational cesium beam clock from Henderson (2005). (N. B. Henderson caption includes 'Crown copyright 1960'. Reproduced by permission of the Controller of HMSO and the Queen's Printer for Scotland).

The concerns of astronomers were described by Essen again in his *Time for Reflection*:

'A few months after the atomic clock had been in operation, the Astronomer Royal invited me to describe it at a meeting of the International Astronomical Union to be held in Dublin. One of the main subjects for discussion was the adoption of a new unit of time. Astronomers knew that the unit based on the rotation of the earth was no longer adequate and they were recommending a unit, the second of the ephemeris time, based on the revolution of the earth round the sun. Unfortunately although this unit might be expected to be more constant than the mean solar second, it is much more difficult to measure, and the observations would have to be averaged over years to give the required accuracy. This rendered it useless as a unit of measurement which must be available immediately. I pointed out that whatever advantages this unit might have for the astronomer it was useless for the physicist and engineer, and suggested that since an atomic unit would be needed in the future it would be wise to defer a decision until agreement could be obtained on the definition of such a unit. There was no support for this suggestion and the second of ephemeris time was adopted and was later confirmed by the International Committee of Weights and Measures, showing how even scientific bodies can make ridiculous decisions. One useful

outcome of the Dublin meeting was that with the help of Markowitz—I was not an official delegate myself—a resolution was passed to the effect that when the relationship between ephemeris time and atomic time was established the atomic clock could be used to make astronomical time available. This meant that we had international approval to introduce atomic time when the comparisons were completed without further international meetings. A detailed programme was arranged with Markowitz. The time interval between certain time signals was measured at the NPL in terms of the atomic clock and at the U.S. Naval Observatory in terms of the ephemeris second. The comparisons took longer than anticipated because of the relative inaccuracy of the astronomical measurements, but after three years it was decided that further averaging was not likely to improve the result. The value was, therefore, announced and was eventually accepted internationally as the unit of time.'

As a result of this collaboration, the frequency of the transition between the two hyperfine states of cesium was determined to be 9 192 631 770 cycles per Ephemeris Time second (Markowitz *et al.*, 1958). This continues to be the basis for the definition of the second using the world's primary frequency standards.

10.4.1.2 Cesium Beam Tubes

The first commercial version of a cesium standard called the 'Atomichron' appeared in 1956 (Figure 10.10). It was manufactured by the National Company



Figure 10.10 The first commercial cesium atomic clock, the 'Atomichron,' with the developers, R. Zacharias (left) and R. T. Daly (right) (Forman, 1998).

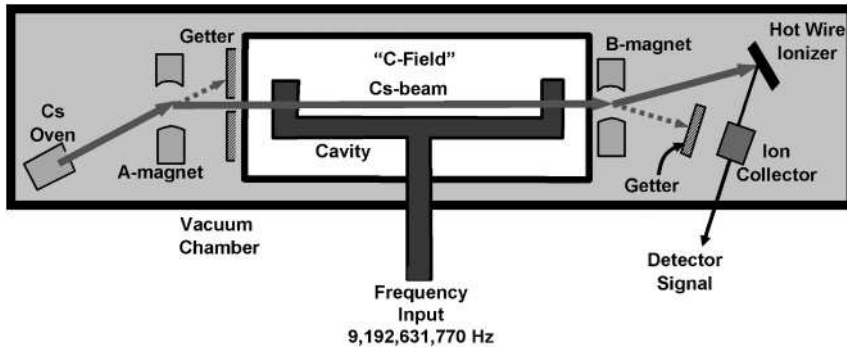


Figure 10.11 Schematic representation of a cesium beam tube. The typical cavity length in commercial tubes is 10 to 20 cm; it is about 4 meters in laboratory standards.

and developed by R. T. Daly, Jerrold Zacharias, and A. Orenberg (Forman, 1998). A period of technological development followed the introduction of the Atomichron, resulting in a cesium atomic beam tube that was about 30 cm long. This tube was eventually incorporated in the Hewlett Packard HP5060 Cesium Atomic Beam Frequency Standard, which quickly became a standard device for high-precision timekeeping.

The basic beam tube configuration is shown in Figure 10.11. Cesium atoms are heated to a temperature of about 90°C and exit the oven traveling with a speed of about 260 ms^{-1} . Magnetic state selection is used to prepare the proper atomic state. This technique makes use of the fact that atoms in states ($F = 3, m_F = 0$) and ($F = 4, m_F = 0$) follow different paths when subjected to magnetic fields. Atoms in the hyperfine state ($F = 3, m_F = 0$) are deflected toward the axis of the tube, while those in the state ($F = 4, m_F = 0$) are deflected off the axis. They go on to contact the walls or ‘getters’ and are eliminated. Those atoms in the proper state then enter the C-field region where they encounter a uniform magnetic field shielding them from the Earth’s or any other stray fields.

The cesium atoms then enter a region where they are exposed to two oscillatory fields. This method of using two oscillatory fields, invented by Ramsey (1950), allows atoms to pass through two equally long regions in which microwave magnetic induction is applied. A drift space separates the two regions. The phase of the microwave field in the second region is set to maximize the probability that the hyperfine transition occurs. In commercial cesium beam tubes the two regions with the oscillatory fields are about 1.5 cm in length and the drift space is about 15 cm long. Laboratory standards may have extended drift regions of a few meters in length.

When the frequency (9 192 631 770 GHz) of the microwave signal is applied properly, the transition ($F = 3, m_F = 0$) to ($F = 4, m_F = 0$) occurs. A second set of magnets is then used to send those atoms in the ($F = 4, m_F = 0$) state in the direction of a hot wire detector made of a metal such as tungsten, platinum,

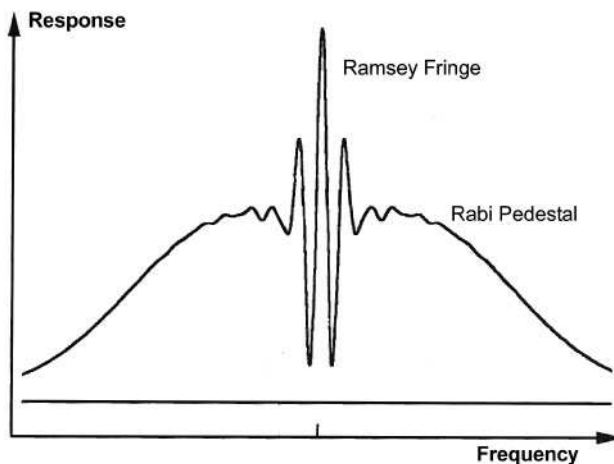


Figure 10.12 Frequency response of cesium beam tube.

or tantalum, or of an alloy such as platinum-iridium, where the atom stream is converted into an electric current and amplified in an electron multiplier. This current is then monitored and maximized by adjusting the frequency of the microwave radiation to provide the maximum number of atoms in the proper state (Audoin, 1992).

The output current measured by the detector depends on the frequency of the microwave energy inserted into the resonator. The response as a function of frequency is shown in Figure 10.12. The central peak is called the Ramsey fringe.

A later development in using cesium beams was the introduction of optical pumping and optical detection of the resonance signal. The technique, proposed by Alfred Kastler (1950) would replace the magnet state selectors by optical interaction regions. One way to improve the signal-to-noise ratio is to ensure that the atomic states are modified to the greatest extent possible. This, in turn, means that the difference between the number of atoms in the ($F = 3, m_F = 0$) and the ($F = 4, m_F = 0$) states should be as large as possible. Optical pumping provides an improvement in the signal-to-noise ratio by increasing the difference in the number of atoms in the states of interest. Figure 10.13 illustrates one possible application of the principle.

When the cesium atoms exit the oven, the ($F = 3, m_F = 0$) and the ($F = 4, m_F = 0$) levels of the ground state ($l = 0$) are approximately equally populated. In the case illustrated in Figure 10.13 the atoms absorb light with a wavelength of $0.85\ \mu\text{m}$ (infrared) raising them from the $F = 4$ ground state to the $l = 1, F = 3$ level. From there they spontaneously decay within a few nanoseconds by emitting radiation and return to either the $F = 3$ or $F = 4$ levels of the ground state. This results in a net transfer of the atoms from the $F = 4$ to $F = 3$ levels of the ground state. The distribution of the atoms among the magnetic sub-levels depends on the transition

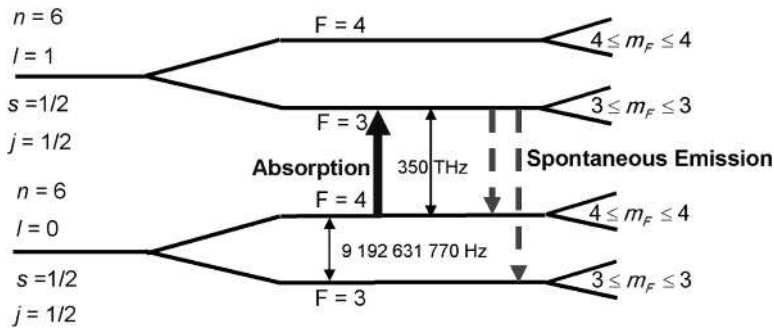


Figure 10.13 Optical pumping. An electron in the $l=0$, $F=4$ state is exposed to light of the proper frequency (approximately 350 THz) to excite it into the $l=1$, $F=3$ state from which it spontaneously emits a photon and decays to either $l=0$, $F=3$ or $l=0$, $F=4$. The net result is to increase the number of atoms with their outer electrons in the $l=0$, $F=3$ state while reducing the number in the $l=0$, $F=4$ state.

probabilities and on the polarization of the light (Picqué, 1977; Audoin and Guinot, 2001).

Just as in the case of the magnetic state selection technique, the atoms are then exposed to the microwave radiation to cause them to transition to the ground state $F=4$ level. Lasers can also be used to replace the B magnets that are used to detect the atoms in this level. Here again, there are various possibilities. If, for example, light from the same laser is used to interact with the atoms exiting the microwave cavity (see Figure 10.14), those in level $F=4$ will be transferred back to the $F=3$ level by emitting a photon. This fluorescence, which is proportional to the number of atoms in the $F=4$ level, can be detected optically and used as a means to adjust

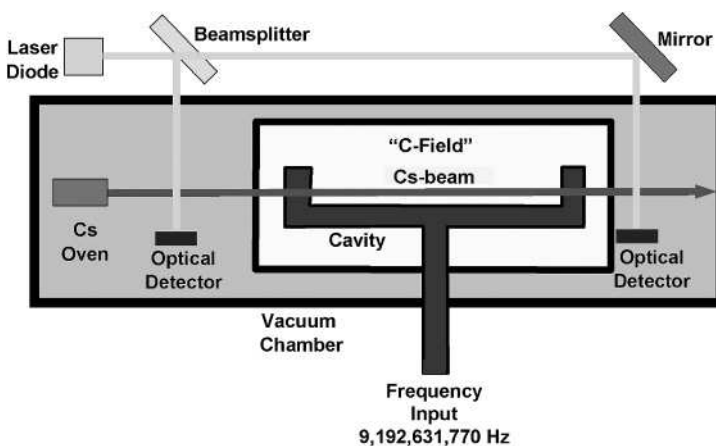


Figure 10.14 Optically pumped cesium beam tube using the light from a single laser.

the frequency of the microwave radiation to provide the maximum number of atoms in the proper state.

Two different lasers can also be used in optical pumping techniques. Referring to Figure 10.4, the first polarized laser, which is tuned to the $F = 3$ ground state transition to the $F = 4$ sub-level of the $^2P_{3/2}$ level, is used to pump atoms into the $F = 4$ level of the ground state in a manner similar to that described above. Making use of selection rules that describe some general rules about permitted transitions, another laser, tuned to the transition between the ground state $F = 4$ level and the $F = 4$ level of the $^2P_{3/2}$ level, removes atoms from the $F = 4$ level except for those in the $m_F = 0$ sublevel. The atoms prepared by this process are then sent to the resonance cavity where the frequency is tuned to the transition between the two hyperfine levels causing stimulated recombination. The atoms in the $F = 3$ level in this case can then be detected optically.

The advantage of optical pumping is that it can enhance the number of atoms in the desired state rather than just reject the atoms in the undesired state. Using optical pumping for site selection and optical detection does not alter the source of the cesium beam or the microwave cavity. By eliminating the need for the state-selection magnet, a larger number of atoms contribute to the signal, which results in a superior signal-to-noise ratio. It also removes the need for strong magnets near the resonance cavity. The first operational optically pumped cesium device appeared in 1980 (Arditi and Picqué, 1980; Ardit, 1982).

10.4.1.3 Cesium Fountains

As mentioned in Section 10.4.1, one of the early attempts to use the cesium atom for timekeeping was that of Jerrold Zacharias, who in 1952 developed the concept of directing a thermal beam of atoms vertically through a microwave cavity. Some of the slower atoms would be expected to return to the source under the influence of gravity. This would provide a time of flight of the order of a second and permit Ramsey interrogation twice by the same microwave cavity. The implementation was not successful, however, because of scattering processes within the beam (Sullivan *et al.*, 2001). Success would have to wait for the development of laser cooling of atoms.

Thermal broadening of the resonance line is one source of instability in a cesium beam atomic clock. It follows, then, that a possible improvement might be to reduce the speed of the atoms in the beam by cooling them to a very low temperature. The speed of the atoms when they exit the 100°C oven is about 260 m s^{-1} . In the 1970s techniques were developed to make it possible to cool the atoms to the extent that they showed hardly any perceptible motion (Wineland and Itano, 1979). In the case of cesium this was made possible by the development of semi-conducting lasers that emitted light with a wavelength of $0.85\text{ }\mu\text{m}$ with high spectral purity. Doppler cooling of atoms is illustrated in Figure 10.15. Monochromatic light of the same intensity propagating in opposite directions is shown directed at an atom of velocity v . For cesium the wavelength of the light is associated with the transition from the ground state to the $l = 1$, $j = 3/2$ level and is tuned to have a frequency slightly less than the resonance frequency. The atom traveling toward

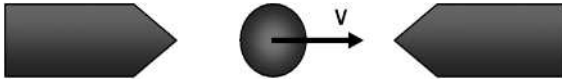


Figure 10.15 Doppler cooling.

the light propagating from the right, in this case, sees the frequency of the light in the direction it is traveling Doppler shifted toward its resonant frequency and shifted away from its resonant frequency in the opposite direction. Consequently it absorbs more photons in the direction of travel, and the momentum exchange tends to slow its motion in that direction. When a cloud of atoms is exposed to light from three pairs of lasers in mutually orthogonal directions for a sufficient length of time to allow for tens of thousands of individual interactions, the atoms are slowed in three dimensions resulting in the term ‘optical molasses.’ This technique can be used to cool cesium atoms to a temperature close to $120\mu\text{K}$.

Even lower temperatures are possible using an additional cooling technique called the ‘Sisyphus effect’ that makes use of spatial variations in the polarization of the light in two opposing beams of light. The interaction of atoms with the electromagnetic field of a laser light beam causes their energy levels to change. If, for example, two laser beams are oriented in opposite directions with equal amplitudes and perpendicular linear polarization, they create a standing wave with polarization that varies with position and repeats spatially at intervals of $\lambda/2$. The energy of the sublevels of an atom interacting with these beams will likewise vary according to the position of the atom along the axis of the light beams. As the atom (assuming it has two energy levels) moves along this axis it will experience a series of hills and valleys of potential energy (Figure 10.16).

When the atom reaches the top of a potential hill, conservation of energy requires that its kinetic energy be lowered. At that point it transitions to a higher

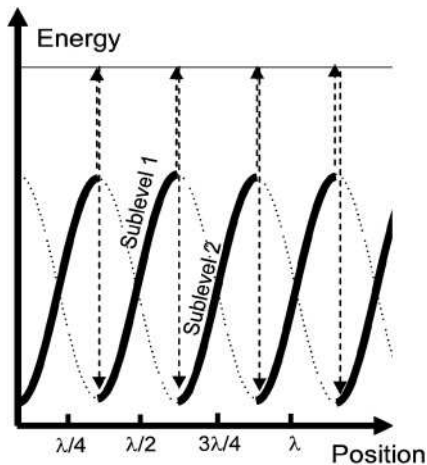


Figure 10.16 Sisyphus cooling.

energy level and subsequently emits a photon, removing the kinetic energy it gained in getting to the top of the potential hill. In emitting the photon it transitions to the second level, which is now at a potential valley. This series of alternating potential highs and lows is called the Sisyphus effect referring to the Greek mythological character Sisyphus who was endlessly doomed to push a rock up a hill only to have it roll down again.

As the atom progresses along the axis the atoms experience a series of potential hills and valleys until they reach a point where they lack the kinetic energy to climb another hill. This mechanism can result in atoms cooled to the level of a few μK . The possibility of cooling atoms to these temperatures made it possible to re-visit the fountain concept.

The first demonstration of a laser-cooled fountain clock was developed in the late 1980s using sodium atoms (Kasevich *et al.*, 1989). The first cesium fountain was built at the Laboratoire Primaire du Temps et des Fréquences in Paris in 1991 (Clairon *et al.*, 1996). Cesium fountains typically cool about 10^7 atoms using six lasers tuned to the $l = 0, F = 4$ to $l = 1, F = 5$ transition with a wavelength of 852 nm. The resulting ball of cooled atoms in the $l = 0, F = 4$ state would then be launched vertically in a cesium fountain with a velocity of about 4 m s^{-1} (see Figure 10.17). This is done by introducing a change in the frequency between the laser beams in the vertical direction. The atoms then pass through a state preparation cavity where they can be optically pumped to the $l = 0, F = 3, m_F = 0$ state. Next, they are

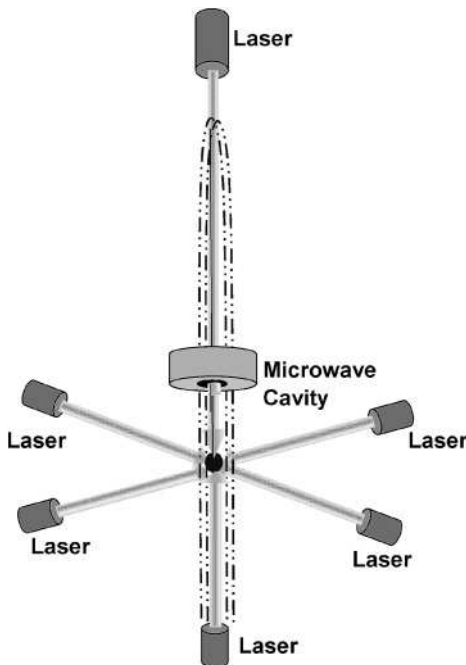


Figure 10.17 Schematic representation of a cesium fountain.

irradiated by the clock radiation and continue drifting upward for a drift period of about $\frac{1}{2}$ s, when gravity causes them to begin to fall back down. As with the cesium beam tube they then experience the clock radiation a second time and pass through to an optical detection region where the number of atoms in the $l = 0, F = 4$ state is sensed. The cycle can then be repeated when another ball of atoms is cooled.

The advantages of cesium fountain technology over the cesium beam technology are (i) the laser cooling and preparation techniques make it possible to allow more atoms to be used since fewer are discarded in the selection process, (ii) possible phase differences between the two arms of the clock radiation cavity in the beam tube are eliminated, and (iii) a significantly longer interrogation time is possible. On the other hand, the beam tube provides a continuous beam of atoms, while the fountain operates with a series of pulses of atoms.

10.4.2

Hydrogen

Hydrogen masers have found significant use in timekeeping since the 1960s. Early work using the single-electron hydrogen atom began in 1960 (Goldenberg, Kleppner and Ramsey, 1960) with a concept based on the ground state hyperfine transition $F = 0$ to $F = 1$ in the hydrogen atom and its corresponding frequency of 1420405751.770 Hz (Audoin and Guinot, 2001). The principle relies on the coupling between the atoms and microwave energy field in a resonant cavity. The stimulated emission of radiation in the maser causes an amplification of the microwave field that sustains the required oscillation. The basic concept is shown in Figure 10.18.

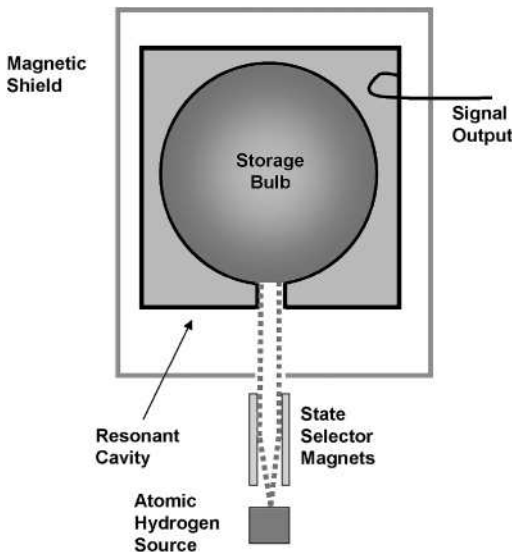


Figure 10.18 Schematic representation of a hydrogen maser.

An atomic hydrogen beam is created and passed through an area that magnetically selects those atoms in the $l = 0$, $F = 1$, $m_F = 1$ and $l = 0$, $F = 1$, $m_F = 0$ states to pass through to a quartz storage bulb whose inner wall is coated with a polymer containing fluorine such as Teflon. Approximately 10^{12} to 10^{13} atoms enter per second where they undergo 10^4 to 10^5 collisions per second. The storage bulb is contained in a resonant cavity, which is contained behind a magnetic shield that maintains a constant magnetic field. Within the storage bulbs the hydrogen atoms undergo the transition from the $F = 1$ to the $F = 0$ level. The atoms entering the cavity amplify the applied frequency field provided that the frequency is tuned to be close to the frequency of the hyperfine atomic transition (Kleppner *et al.*, 1962, 1965). Hydrogen masers differ from the cesium devices in that there is no direct measure of the change in the population of atoms in the different energy states, because there is no efficient means of detecting hydrogen atoms.

The stability of the hydrogen maser depends critically on the tuning of the cavity frequency. A variety of procedures can be used to tune that frequency automatically. In all cases an error signal is determined and that signal is then used to tune the cavity frequency. Sources of error in hydrogen masers include thermal noise, collisions between atoms, collisions with the storage bulb walls, and nonuniformity of the magnetic field (Audoin and Guinot, 2001). Operationally the maser can be used as an active or passive device for precise time-keeping.

10.4.2.1 Active Hydrogen Maser

An active hydrogen maser makes use of a quartz crystal phase locked to the output signal of the hydrogen maser. The maser operates spontaneously and its signal is detected by an antenna in the resonant cavity and synchronized to the crystal output as shown in Figure 10.19. The frequency produced by the quartz oscillator is mixed with the signal from the active maser to produce a signal that is used to steer the output of the oscillator. Some advantages of an active hydrogen maser are the narrow width of the resonant frequency due to the relatively long storage time, the unperturbed movement of the atoms due to the low pressure, the low velocity of the atoms, and the low noise level.

Efforts to develop a cryogenic hydrogen maser (Vessot, Levine and Mattison, 1977; Crampton *et al.*, 1979; Vessot, Mattison and Blomberg, 1979; Hess *et al.*, 1986; Hürlimann *et al.*, 1986; Walsworth *et al.*, 1986) began in the late 1970s. The concept is to reduce thermal noise in the atoms, which could lead to a possible improvement in stability of three orders of magnitude over that achieved by a hydrogen maser operating at room temperature. One such device has been operated at a temperature of 0.5 °K, achieving results similar to that of an active hydrogen maser at room temperature for averaging times less than about 100 s. For averaging times greater than that, results have not matched those of a room-temperature maser. The cryogenic maser may find application in combination with other devices under development, but further studies are needed (Vessot, 2005).

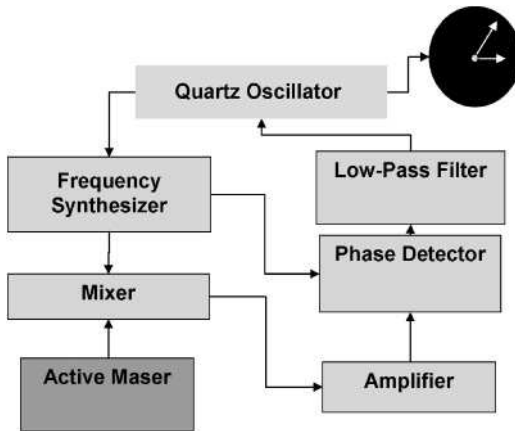


Figure 10.19 Block diagram of an active hydrogen maser.

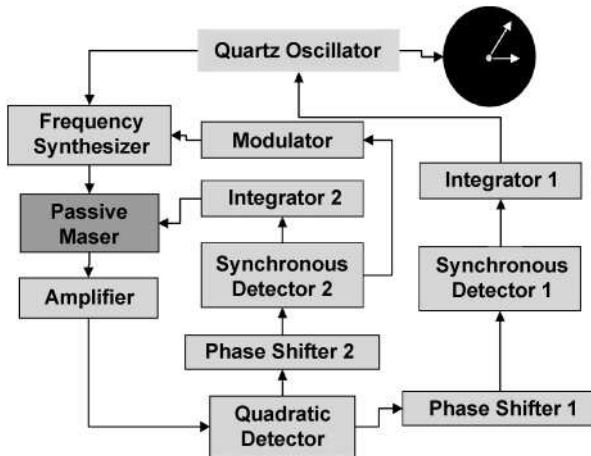


Figure 10.20 Block diagram of a passive hydrogen maser.

10.4.2.2 Passive Hydrogen Masers

In contrast with the active hydrogen maser, the passive device steers the quartz oscillator by locking the output of the crystal to the maser signal, as is done with cesium oscillators. In the passive hydrogen maser, however, resonance is monitored by measuring the changes in amplitude of the electromagnetic field in the cavity. Both the frequency of the magnetic field in the cavity and the quartz oscillator frequency can be controlled following the process illustrated in Figure 10.20. The maser output signal is first amplified. The appropriate signals are detected and sent to two circuits: one is used to determine the difference between the interrogation signal carrier frequency and the atomic resonance frequency, and the

second determines the difference between the carrier frequency and the cavity resonance. These are used to provide two error signals, one to control the quartz oscillator and the other the cavity resonance frequency.

10.4.3

Rubidium

Another alkali atom, rubidium, is used extensively for precise timekeeping. It is a silvery-white metallic element having atomic number 37 that liquefies at 39.3°C and reacts vigorously in water. There are 24 isotopes of rubidium, but there are only two that occur naturally: Rb-85 which comprises 72.2% of the naturally occurring rubidium and Rb-87, which is slightly radioactive with a half-life of 4.88×10^{10} years that comprises the remainder. Devices using rubidium make use of the ground state hyperfine clock transition of Rb-87 with a corresponding frequency of 6 834 682 610.904 Hz. Development of these standards began in the late 1950s (Arditi, 1958; Carpenter *et al.*, 1960; Ardit and Carver, 1961, 1964; Packard and Swartz, 1962; Davidovits, 1964; Davidovits and Novick, 1966).

10.4.3.1 Rubidium Cells

Rubidium cell clocks generally begin with a lamp, operating at a temperatures ranging between 65°C and 140°C, that contains a small amount (about one mg) of Rb-87 or a mixture of Rb-85 and Rb-87, along with a noble gas such as krypton. A radio frequency generator creates a light discharge (see Figure 10.21) that is passed through a hyperfine filter that contains Rb-85 and a noble gas at high pressure (about 10^4 Pa). By a coincidence of the structure of rubidium, the atoms of Rb-85 in the filter absorb the components of the light emitted in the $l = 1$ to the ground state $F = 2$ level, resulting in a light with the spectrum shown in Figure 10.21. This is then passed to the resonance cell in a resonant cavity tuned to the resonance frequency of Rb-87 that contains Rb-87 along with nitrogen and rare gases. There the light is absorbed by the Rb-87 atoms in the $F = 1$ level of the ground state. They are then optically pumped to the $l = 1$ states, where they relax to either the $F = 1$ or $F = 2$ levels of the ground state. The net effect is to depopulate the $F = 1$ level and populate the $F = 2$ level. The resonance frequency in the cavity, however, stimulates the transition from the $F = 2$ to the $F = 1$ levels, thereby repopulating the $F = 1$ level. These atoms, in turn, absorb part of the incident light, reducing the intensity detected by the photodiode. This signal can be used to adjust the resonance frequency delivered to the cavity to achieve a minimum intensity (Audoin and Guinot, 2001). The buffer gases are used to keep the rubidium atoms away from the walls of the cell.

The additional gases in the resonance cell are used to damp the Doppler broadening of the resonance frequency. Various modifications can be made to this basic design. Frequently the filter is combined with the resonance cell to reduce the size of the standards, for example. These rubidium cell standards are not as accurate as cesium standards because the collisions with the molecules of the buffer gases introduce instabilities. Cesium standards are able to allow the cesium atoms to

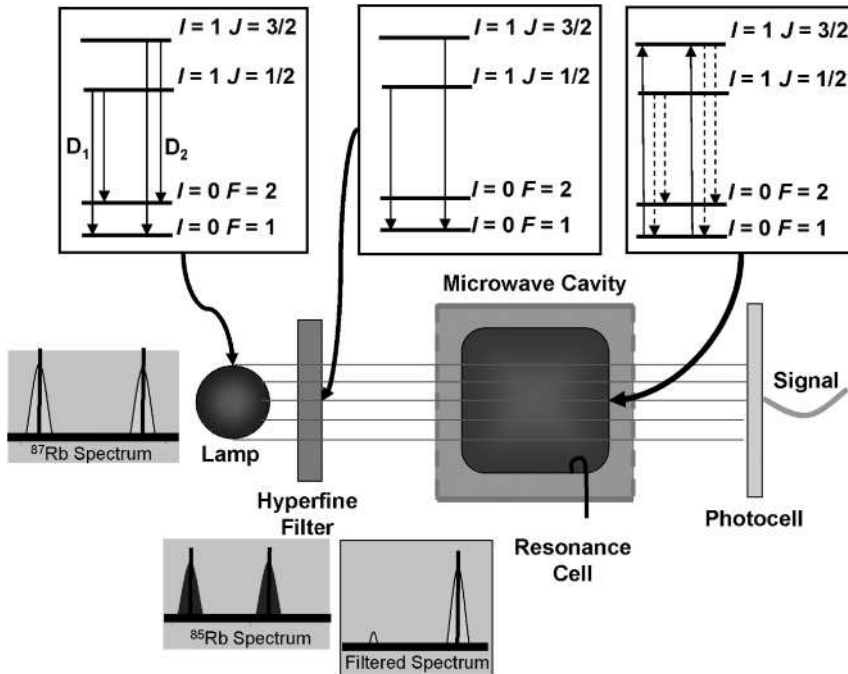


Figure 10.21 Rubidium Cell frequency standard.

drift through an extended cavity and take advantage of the Ramsey-separated fields to narrow the line width of the resonance frequency. Rubidium standards are also sensitive to environmental conditions. Optical pumping does, however, introduce a shift in the hyperfine structure of the atom, called light shift, which affects the accuracy of the device.

10.4.3.2 Rubidium Fountains

The largest source of instability in cesium fountains is the cold-collision shift, but this effect is much smaller for rubidium atoms. Also, because the clock frequency used for rubidium is smaller than that for cesium, the entrance and exit dimensions of the microwave cavity can be larger, allowing the use of more atoms. These considerations, along with the fact that rubidium has a less complicated series of magnetic states with which to contend, make the concept of a rubidium fountain attractive in spite of its lower resonant frequency (Fertig and Gibble, 2000). The principles of operation of a rubidium fountain are similar to those of a cesium fountain, except that cesium is replaced with rubidium.

10.4.3.3 Double-bulb Rubidium Maser

A further application of rubidium in precise timekeeping is the double-bulb rubidium maser proposed in 1994 (Golding *et al.*, 1994). To mitigate the effect of light

shift in this design, optical pumping occurs in a bulb separated from the region in which the microwave interaction occurs. No buffer gas is required in this design, and the atoms are 'recycled' in that they are able to be returned to the bulb, where the pumping occurs, to be 're-pumped.' The lack of buffer gases in this design allows the atoms to effuse between the two areas, but this also makes the device sensitive to interactions of the atoms with the walls of the bulbs. This problem can be mitigated by a judicious choice of wall coatings. One of its advantages is its relatively small size compared with a hydrogen maser.

10.5 Stored Ion Clocks

The sources of error in all atomic clocks include perturbations due to confinement of the atoms. To eliminate the effect of wall collisions, ions trapped in an electromagnetic field can be used to provide a timing source. The basic concept for applying stored ions in timekeeping is the same as that for atomic clocks. A local oscillator is used to provide a frequency near a resonant atomic frequency, and the excited transitions between states provide an accurate frequency via a feedback chain. A number of singly ionized atoms with well-known hyperfine structure in the ground state are potential candidates for this application. Two basic kinds of ion traps can be used. These are the Paul trap and the Penning trap. Static electric and magnetic fields are used to confine ions in a Penning trap. A Paul trap makes use of a radio frequency field.

In a Penning trap the plates on the ends (see Figure 10.22) are kept at the same potential with respect to the inner ring electrode. This field forces ions toward the center of the trap if the ions are displaced in a direction toward either of the end plates. If they are displaced in any direction parallel to the planes of the end plates, however, a magnetic field B is required to force the atoms to the center of the trap. As a result, a single ion will then undergo simple harmonic motion in the z direc-

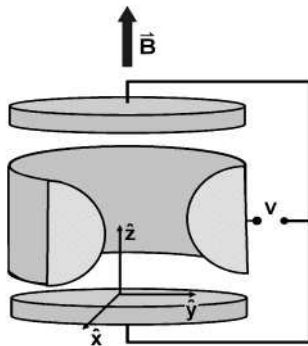


Figure 10.22 Schematic representation of a Penning ion trap.

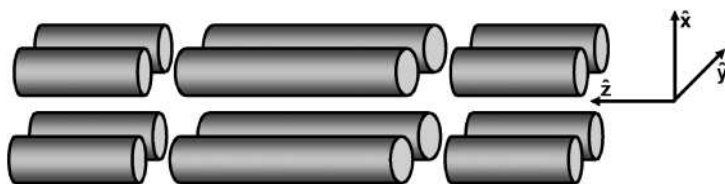


Figure 10.23 Linear Paul ion trap.

tion, and a circular cyclotron motion along with another lower-frequency circular magnetron ($E \times B$) motion in the x - y plane.

In a Paul trap the electric potential between the inner element and the end plates oscillates at a high frequency that is determined by the geometry of the trap, and no magnetic field is required. In this configuration an ion will undergo motion with a frequency equal to the frequency of the electric potential that is applied. This is called the ‘micromotion.’ It also experiences motion with a much lower frequency, called the ‘secular motion’ (Itano, 1991). A variation is the linear ion trap shown in Figure 10.23 (Prestage, Dick and Maleki, 1989). In this design an RF potential is applied between the rods. The phase at each rod differs by 180° from its neighbors. Such a trap attracts ions to a central axis where they have less micromotion in comparison with other types of traps. The rods can be connected in rings, resulting in a ‘racetrack’ trap (Church, 1969).

By trapping the ions in such traps, they can be allowed to interact with microwaves for several seconds. The longer they can interact without being disturbed, the more stable is the frequency standard. A number of ionized atoms with well-known ground state hyperfine structure are available. These include helium, beryllium, magnesium, strontium, calcium, barium, ytterbium, cadmium, and mercury.

10.5.1

Mercury

The first practical application of trapped ionized atoms for timekeeping made use of the element mercury (chemical symbol Hg). The transition of interest is the hyperfine transition of $^{199}\text{Hg}^+$, which has a very high frequency (40 507 347 996.841 Hz) and high Q . The measurement of the transition frequency was originally done in 1973 (Major and Werth, 1973), but the first frequency standard based on this technique was developed in the early 1980s (Jardino *et al.*, 1981). Since that time there has been extensive development activity (Cutler *et al.*, 1985, 1986; Tjoelker *et al.*, 1996, 2000; Prestage, Tjoelker and Maleki, 2001).

Mercury is used because its transition frequency is less affected by systematic effects related to magnetic fields and ion motions than would be the case in lighter atoms. This allows for better long-term stability while minimizing environmental issues. In an operational cycle an electron gun ionizes the atoms of neutral ^{199}Hg introduced into the device (see Figure 10.24). A cloud of $^{199}\text{Hg}^+$ ions with a density of about 10^6 ions cm^{-3} is used. State selection is done using optical

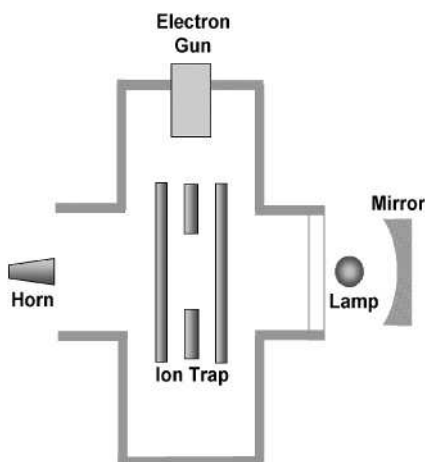


Figure 10.24 Linear ion trap frequency standard.

pumping with a wavelength of 194 nm. This process takes advantage of a fortunate coincidence. The wavelength of the D_1 line of the spectrum of the $^{202}\text{Hg}^+$ ion of the mercury isotope of atomic weight 202 is very close to the wavelength corresponding to the transition between the $l = 0, F = 1$ ground state level to the $l = 1, F = 1$ level. A lamp providing the 194 nm wavelength light using $^{202}\text{Hg}^+$ ions is used to pump the $^{199}\text{Hg}^+$ ions to the $l = 0, F = 0$ level. They can then be exposed to the 40.5 GHz hyperfine transition frequency that will repopulate the $l = 0, F = 1$ level. When they decay back to the $l = 0, F = 0$ level, the fluorescence will intensify the 194 nm light and the intensity of this signal is monitored to determine the clock transition frequency. A buffer gas is mixed with the $^{199}\text{Hg}^+$ ions. A device following such a scheme called the Linear Ion Trap Frequency Standard (LITS) has been put in operation at the Jet Propulsion Laboratory (Tjoelker *et al.*, 1996).

A variation of the LITS that makes use of two chambers is also being developed at the Jet Propulsion Laboratory (Prestage *et al.*, 1995). It is called an extended linear ion trap (LITE) or shuttle trap. Two separate trap regions are used: a region for microwave interrogation and a separate volume for loading and state preparation. Ions are shuttled between these two regions, hence the name ‘shuttle trap.’ This device provides for an interrogation region seven times longer than that of the LITS and is less sensitive to environmental conditions. Improvement in stability by a factor of two compared with that provided by a single-chamber device can be expected. Mercury stored-ion devices are limited in their stability, not only by detection noise, but by magnetic and thermal effects and collisions with the buffer gas molecules.

A laser-cooled linear ion trap using mercury ions has been developed (Berkeland *et al.*, 1998). It uses the same Doppler cooling technique that was described for cooling atoms in a Paul trap to form a linear mercury crystal of

seven ions. Light with wavelength of 198 nm is used to cool the ions without using a buffer gas.

10.5.2

Other Ions

Ions other than those of mercury have also been used in developmental frequency standards. These standards follow the same basic principles of the mercury ion trap, but have not enjoyed widespread use. The first laser-cooled ion trap used a Penning trap with $^9\text{Be}^+$ (beryllium) ions (Bollinger *et al.*, 1985). Other ions involved include $^{173}\text{Yb}^+$ (ytterbium) (Münch *et al.*, 1987), $^{171}\text{Yb}^+$ (ytterbium) (Fisk *et al.*, 1995), $^{25}\text{Mg}^+$ (magnesium) (Itano and Wineland, 1981), $^{137}\text{Ba}^+$ (barium) (Blatt and Werth, 1982), and $^{135}\text{Ba}^+$ (barium) (Becker and Werth, 1983).

10.6

Characterizing Atomic Clocks

The stabilities of the various microwave atomic time-keeping standards are shown in Figure 10.25. Standards that are still in development, such as the fountains or ion traps, may be expected to improve as research leads to new developments.

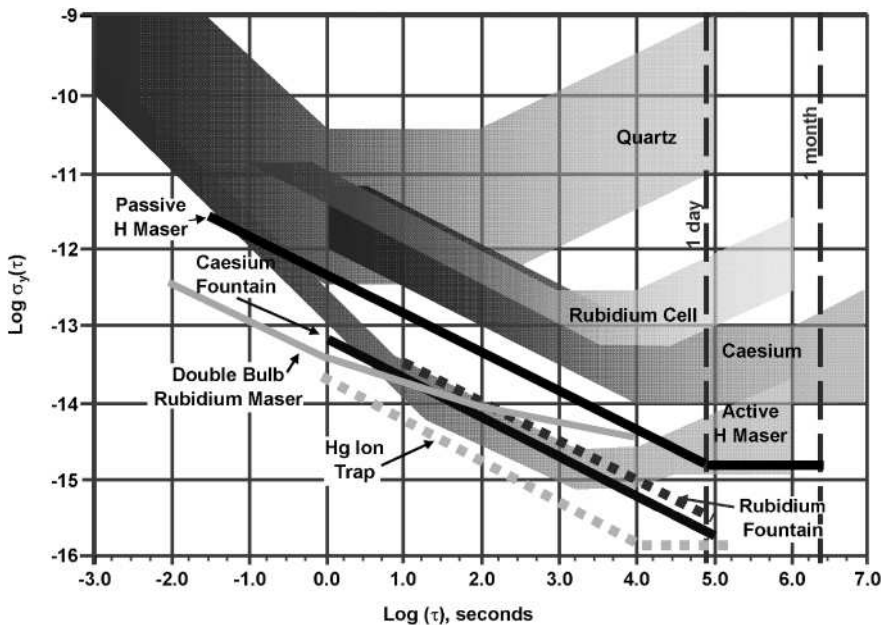


Figure 10.25 Stabilities of microwave timekeeping standards.

References

- Arditi, M. (1958) L'influence des gaz tampons sur le déplacement de la fréquence et la largeur des raies des transitions hyperfines de l'état fondamental des atomes alcalins. *Le Journal de Physique et le Radium*, **19**, 873.
- Arditi, M. (1982) A cesium beam atomic clock with laser optical pumping, as a potential frequency standard. *Metrologia*, **18**, 59–66.
- Arditi, M. and Carver, T.R. (1961) Pressure, light, and temperature shifts in optical detection of 0-0 hyperfine resonance of alkali metals. *Phys. Rev.*, **124**, 800–9.
- Arditi, M. and Carver, T.R. (1964) Hyperfine relaxation of optically pumped Rb87 atoms in buffer gases. *Phys. Rev.*, **136**, 643–9.
- Arditi, M. and Picqué, J.-L. (1980) Construction and preliminary tests of a laser optically pumped cesium jet atomic clock. *C R, Série B. Sci. Physiq.*, **290**, 461–4.
- Audoin, C. (1992) Cesium beam frequency standards: classical and optically pumped. *Metrologia*, **29**, 113–34.
- Audoin, C. and Guinot, B. (2001) *The Measurement of Time*, Cambridge University Press, Cambridge.
- Becker, W. and Werth, G. (1983) Precise determination of the ground state hyperfine splitting of $^{135}\text{Ba}^+$. *Zeitschrift für Physik A*, **311**, 41–7.
- Berkeland, D.J., Miller, J.D., Bergquist, J.C., Itano, W.M. and Wineland, D.J. (1998) Laser-cooled mercury ion frequency standard. *Phys. Rev. Lett.*, **2089**–92.
- Blatt, R. and Werth, G. (1982) Precision determination of the ground state hyperfine splitting in $^{137}\text{Ba}^+$ using the ion-storage technique. *Phys. Rev. A*, **25**, 1476–82.
- Bollinger, J.J., Prestage, W.M., Itano, W.M. and Wineland, D.J. (1985) Laser-cooled-atomic frequency standard. *Phys. Rev. Lett.*, **54**, 1000–3.
- Carpenter, R.J., Beaty, E.C., Bender, P.L., Saito, S. and Stone, R.O. (1960) A prototype rubidium vapor frequency standard. *IRE Trans. Instrum.*, **1-9**, 132–5.
- Church, D.A. (1969) Storage-ring ion trap derived from the linear quadrupole radio-frequency mass filter. *J. Appl. Phys.*, **40**, 3127–34.
- Clairon, A., Ghezali, S., Santarelli, G., Laurent, P., Lea, S.N., Bahoura, M., Simon, E., Weyers, S. and Szymaniec, K. (1996) Preliminary accuracy evaluation of a cesium fountain frequency standard, in *Proc. Fifth Symp. on Freq. Standards and Metrology* (ed. J.C. Bergquist), World Scientific, London, pp. 49–59.
- Crampton, S.B., Kleppner, D., Phillips, W.D., Weinrib, A., Greytak, T.J., Smith, D.A. (1979) Hyperfine resonance of gaseous atomic hydrogen at 4.2 K. *Phys. Rev. Lett.*, **42**, 1039–42.
- Cutler, L.S., Flory, C.A., Giffard, R.P. and McGuire, M.D. (1986) Doppler effects due to thermal macromotion of ions in an RF quadrupole trap. *Appl. Phys. B*, **39**, 251–9.
- Cutler, L.S., Giffard, R.P. and McGuire, M.D. (1985) Thermalization of ^{199}Hg ion macromotion by a light background gas in an RF quadrupole trap. *Appl. Phys. B*, **36**, 137–42.
- Davidovits, P. (1964) An optically pumped Rb^{87} maser oscillator. *Appl. Phys. Lett.*, **5**, 15–16.
- Davidovits, P. and Novick, R. (1966) The optically pumped rubidium maser. *Proc. IEEE*, **54**, 155–70.
- Essen, L. (2000) Time for Reflection, published privately and available at <http://www.btinternet.com/~time.lord/>, also available in Henderson, D. (2005), Essen and the National Physical Laboratory's atomic clock, *Metrologia*, **42**, S4–9.
- Essen, L. and Parry, J.V.L. (1955) An atomic standard of frequency and time interval: a caesium resonator. *Nature*, **176**, 280–2.
- Essen, L. and Parry, J.V.L. (1957) The cesium resonator as a standard of frequency and time. *Philos. Trans. R. Soc. London A. Math. Phys. Sci.*, **250**, 45–69.
- Fertig, C. and Gibble, K. (2000) Measurement and cancellation of the cold collision frequency shift in an ^{87}Rb fountain clock. *Phys. Rev. Lett.*, **85**, 1622–5.
- Fisk, P.T.H., Sellars, M.J., Lawn, M.A., Coles, C., Mann, A.G. and Blair, D.G. (1995) Very high Q microwave spectroscopy on trapped $^{171}\text{Yb}^+$ ions: application as a frequency standard. *IEEE Trans. Instrum. Meas.*, **44**, 113–16.

- Forman, P. (1998) Atomichron: the atomic clock from concept to commercial product. Available at: <http://www.ieee-uffc.org/freqcontrol/atomichron/atomichron.htm>
- Goldenberg, H.M., Kleppner, D. and Ramsey, N.F. (1960) Atomic hydrogen maser. *Phys. Rev. Lett.*, **5**, 361–2.
- Golding, W.M., Frank, A., Beard, R., White, J., Danzy, F. and Powers, E. (1994) The double bulb rubidium maser. *Proceedings of the 1994 IEEE International Frequency Control Symposium*, pp. 724–30.
- Henderson, D. (2005) Essen and the national physical laboratory's atomic clock. *Metrologia*, **42**, S4–9.
- Hess, H.F., Kochanski, G.P., Doyle, J.M., Greytak, T.J. and Kleppner, D. (1986) Spin-polarized hydrogen maser. *Phys. Rev. A*, **34**, 1602–4.
- Hürlimann, M.D., Hardy, W.N., Berlinsky, A.J. and Cline, R.W. (1986) Recirculating cryogenic hydrogen maser. *Phys. Rev. A*, **34**, 1605–608.
- Itano, W.M. (1991) Atomic ion frequency standards. *Proc. IEEE*, **79**, 936–42.
- Itano, W.M. and Wineland, D.J. (1981) Precision measurement of the ground-state hyperfine constant of $^{25}\text{Mg}^+$. *Phys. Rev. A*, **24**, 1364–73.
- Jardino, M., Desaintfusien, M., Barillet, R., Viennet, J., Petit, P. and Audoin, C. (1981) Frequency stability of a mercury ion frequency standard. *Appl. Phys. A*, **24**, 107–12.
- Kasevich, M., Riis, E., Chu, S. and De Voe, R. (1989) RF spectroscopy in an atomic fountain. *Phys. Rev. Lett.*, **63**, 612–15.
- 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. Application à l'expérience de Stern et Gerlach et à la résonance magnétique. *Le Journal de Physique et le Radium*, **11**, 255–65.
- 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. *Phys. Rev.*, **138**, A972–83.
- Kleppner, D., Goldberg, H.M. and Ramsey, N.F. (1962) Theory of the hydrogen maser. *Phys. Rev.*, **126**, 603–15.
- Lombardi, M.A., Heavner, T.P. and Jefferts, S.R. (2007) NIST primary frequency standards and the realization of the SI second. *Measure*, **2**, 74–89.
- Lyons, H. (1949) The atomic clock. *Instruments*, **22**, 133–5.
- Major, F.G. and Werth, G. (1973) High-resolution magnetic hyperfine resonance in harmonically bound ground state 199HG ions. *Phys. Rev. Lett.*, **30**, 1155–8.
- Markowitz, W., Hall, R.G., Essen, L. and Perry, J.V.L. (1958) Frequency of cesium in terms of ephemeris time. *Phys. Rev. Lett.*, **1**, 105–7.
- McCoubrey, A.O. (1996) History of atomic frequency standards: a trip through 20th century physics. *Proceedings of the 1996 IEEE International Frequency Control Symposium, IEEE*, pp. 1225–41.
- Millman, S. and Kusch, P. (1940) On the radiofrequency spectra of sodium, rubidium and cesium. *Phys. Rev.*, **58**, 438–45.
- Münch, A., Berkler, M., Gerz, C., Wilsdorf, D. and Werth, G. (1987) Precise ground-state hyperfine splitting in ^{173}Yb II. *Phys. Rev. A*, **35**, 4147–50.
- Packard, M.E. and Swartz, B.E. (1962) The optically pumped rubidium vapor frequency standard. *IRE. Trans. Instrum.*, **I-11**, 215–23.
- Picqué, J.-L. (1977) Hyperfine optical pumping of a cesium atomic beam, and applications. *Metrologia*, **13**, 115–19.
- Prestage, J.D., Dick, G.J. and Maleki, L. (1989) New ion trap for frequency standard applications. *J. Appl. Phys.*, **66**, 1013–17.
- Prestage, J.D., Tjoelker, R.J., Dick, G.J. and Maleki, L. (1995) Progress report on the improved linear ion trap physics package. *Proc. 49th Ann. Symp. Freq. Control Symposium*, pp. 82–5.
- Prestage, J.D., Tjoelker, R.J. and Maleki, L. (2001) Recent developments in microwave ion clocks, in *Topics in Applied Physics, Frequency Measurement and Control*, Vol. 79 (ed. A.N. Luiten), Springer, Berlin, Heidelberg, New York, pp. 195–211.
- Rabi, I.I., Millman, S., Kush, P. and Zacharias, J.R. (1939) The molecular beam resonance method for measuring nuclear magnetic moments the magnetic moments of $^3\text{Li}^6$, $^3\text{Li}^7$ and $^9\text{F}^{19}$. *Phys. Rev.*, **55**, 526–35.

- Ramsey, N.F. (1949) A new molecular beam resonance method. *Phys. Rev.*, **76**, 996.
- Ramsey, N.F. (1950) A molecular beam resonance method with separated oscillating fields. *Phys. Rev.*, **78**, 695–9.
- Ramsey, N.F. (1983) History of atomic clocks. *J. Res. Natl. Bureau Stand.*, **88**, 301–20.
- Ramsey, N.F. (1993) *I. I. Rabi, 1898–1988, Biographical Memoir*, National Academy of Sciences, Washington, DC.
- Ramsey, N.F. (2005) History of early atomic clocks. *Metrologia*, **42**, S1–3.
- Sherwood, J., Lyons, H., McCracken, R. and Kusch, P. (1952) High frequency lines in the hfs spectrum of cesium. *Bull. Am. Phys. Soc.*, **27**, 43.
- Sullivan, D.B., Bergquist, J.C., Bollinger, J.J., Drullinger, R.E., Itano, W.M., Jefferts, S.R., Lee, W.D., Meekhof, D., Parker, T.E., Walls, F.L. and Wineland, D.J. (2001) Primary atomic frequency standards at NIST. *J. Res. Natl. Inst. Stand. Technol.*, **106**, 47–63.
- Tjoelker, R.L., Bricker, C., Diener, W., Harnell, R.L., Kirk, A., Kuhnle, P., Prestage, J.D., Santiago, D., Seidel, D., Stowers, D.A., Sydnor, R.L. and Tucker, T. (1996) A mercury ion frequency standard engineering prototype for the nasa deep space network. *Proceedings of the 1996 IEEE/EIA International Frequency Control Symposium and Exhibition*, pp. 1073–81.
- Tjoelker, R.L., Chung, S., Diener, W., Kirk, A., Maleki, L., Prestage, J.D. and Young, B. (2000) Nitrogen buffer gas experiments in mercury trapped ion frequency standards. *Proceedings of the 2000 IEEE/EIA International Frequency Control Symposium and Exhibition*, pp. 668–71.
- Vessot, R.F.C. (2005) The atomic hydrogen maser oscillator. *Metrologia*, **42**, S80–9.
- Vessot, R.F.C., Levine, M.W. and Mattison, E.M. (1977) Comparison of theoretical and observed maser stability limitation due to thermal noise and the prospect of improvement by low temperature operation. *Proc. 9th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting (NASA, Goddard Space Flight Center 29 November–1 December 1977)*, p. 549.
- Vessot, R.F.C., Mattison, E.M., Blomberg, E.L. (1979) Research with a cold atomic hydrogen maser, in *Annual Frequency Control Symposium, May 30–June 1, 1979, Proceedings*, Electronic Industries Association, Washington, DC, pp. 511–14.
- Walsworth, R.L., Silvera, I.F., Godfried, H.P., Agosta, C.C., Vessot, R.F.C. and Mattison, E.M. (1986) Hydrogen maser at temperatures below 1 K. *Phys. Rev. A*, **34**, 2550.
- Wineland, D.J. and Itano, W.M. (1979) Laser cooling of atoms. *Phys. Rev. A*, **20**, 1521–40.

11

Optical Atomic Standards

11.1

Optical Transition Frequencies

The much higher optical transition frequencies provide an attractive alternative to the microwave frequencies, and, with continuing progress in laser technology, standards based on these frequencies are being realized. Figure 11.1 shows the electromagnetic spectrum with the microwave and optical frequencies along with the expected calibration errors of each of the transition frequencies. The high frequencies of the optical transitions, however, pose a problem when applied to timekeeping. They need to be converted to frequencies that can be used in time-keeping circuitry before they can be useful for practical application. Radio frequencies up to about 100 GHz can be measured easily using electronic means, but measurement of frequencies higher than 100 GHz is problematic.

The answer to the problem is the development of frequency combs. These systems take advantage of developments in laser technology. When mode-locked lasers capable of producing light pulses with lengths of the order of femtoseconds (10^{-15} second) became available, frequency combs became possible. These mode-locked lasers are able to provide trains of extremely short optical pulses that are the result of a superposition of many continuous-wave longitudinal cavity modes. The frequency of the n th mode is given by

$$f_n = nf_R + f_0 \quad (11.1)$$

where f_R is the repetition frequency given by the reciprocal of the pulse repetition interval, and f_0 is a frequency offset. The appearance of the spectrum in the frequency domain is then extremely broad and comprises a series of ‘spikes’ spaced according to (11.1), hence the name ‘frequency comb.’ The appearance of the spectrum is then directly related to the pulse rate of the laser. By matching the optical frequency of the laser linked to the energy level transition to the frequency comb by adjusting the repetition rate, we can relate the optical frequency to a signal useful for precise timekeeping (Udem *et al.*, 1999; Holzwarth *et al.*, 2000; Diddams *et al.*, 2002; Ma *et al.*, 2004).

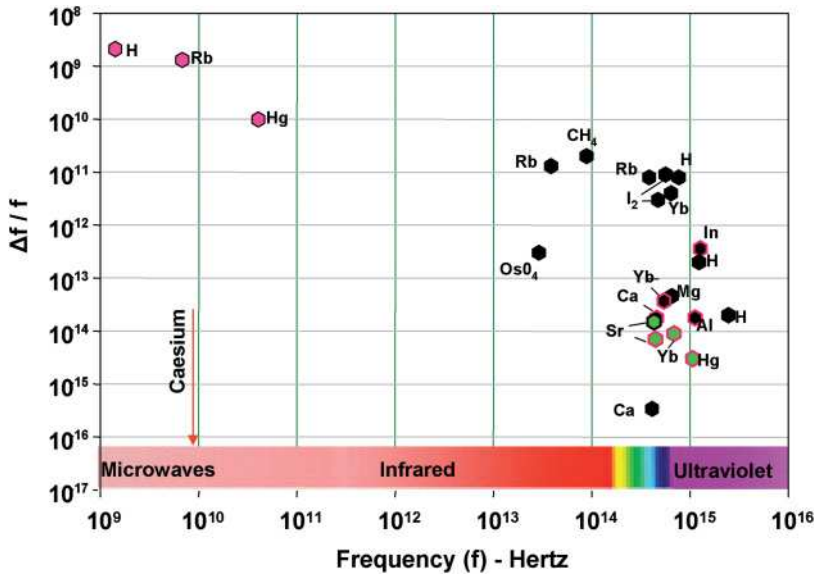


Figure 11.1 Spectrum of transition frequencies of interest in precise timekeeping applications. The calibration error represented as $\Delta f/f$ is shown for each transition. The microwave frequencies are represented in red and the optical frequencies currently

suggested for research purposes by the Consultative Committee for Time and Frequency of the Comité International des Poids et Mesures (CIPM) are shown in green. The standard cesium transition that is used to define the SI second is also shown.

An optical clock, then, is quite similar to a microwave standard, but it uses a stabilized laser as the local oscillator, the output of which is used in a feedback loop to produce the desired energy level transition. The feedback signal is provided by means of a cooling laser that is also used to detect the occurrence of energy level transitions (Figure 11.2). When the light from the laser, acting as the local oscillator, causes transitions to occur, then no fluorescence occurs from photons involved in the cooling transition. By monitoring the level of the fluorescence it is possible to obtain the frequency of the energy level transition.

The output signal of the clock is produced using a frequency comb based on another laser producing femtosecond pulses. This laser emits pulses at a nominal repetition rate f_R , and this, in turn, provides a comb of frequencies spaced at intervals f_n as given by (11.1). If the comb spans one octave (a factor of two in frequency), the term f_0 can be derived by a self-referencing technique. In this process the frequency of the n^{th} infrared mode is doubled, producing a visible frequency that can be heterodyned with the n^{th} mode frequency. The heterodyne signal provides a frequency $2(nf_R + f_0) - (2nf_R + f_0) = f_0$.

A second beat frequency, f_b is also obtained between the m^{th} frequency mode, $f_m = mf_R + f_0$, and the local oscillator frequency f_L . The clock then uses two feedback phase-locked loops to control f_0 and f_b . The first of these loops controls the femto-

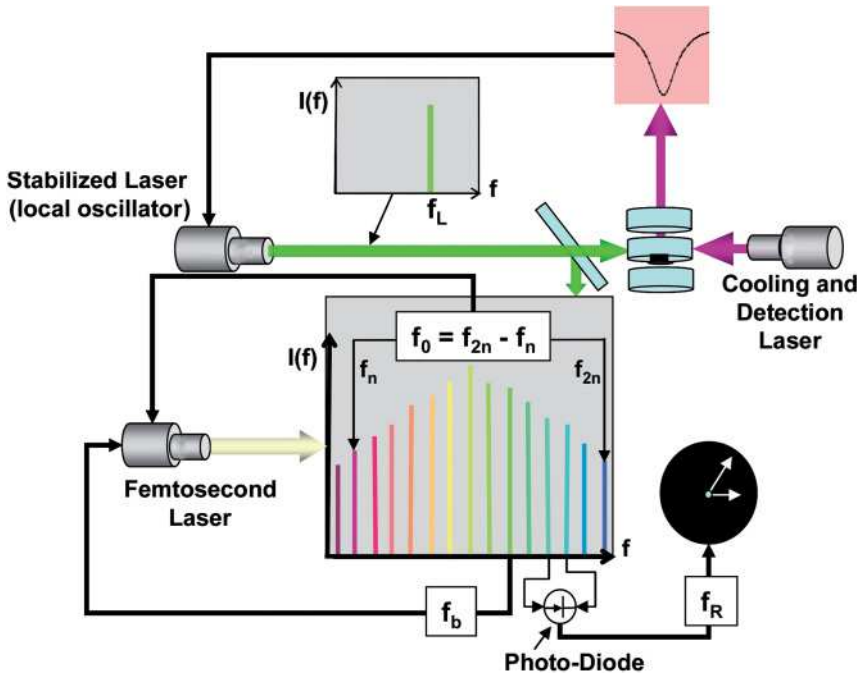


Figure 11.2 Schematic representation of an optical clock.

second laser power so that f_0 maintains a constant relationship with f_R given by $f_0 = \beta f_R$. The second loop controls the cavity length of the femtosecond laser so that $f_b = \alpha f_R$. In this way the frequencies of the two loops are phase referenced to the frequency of the local oscillator. A photo-diode is used to measure the frequency f_R , which is used as the clock frequency (Diddams *et al.*, 2001).

No optical clocks have yet been operated for extended periods of time. They can be operated for short periods of time to calibrate other frequency standards, but they have not been used in the capacity of continuously operating clocks.

11.2 Optical Ion Clocks

The idea of using a single ion as a frequency standard was first proposed by Hans Dehmelt (1982). The principles of an optical clock have now been realized practically in a mercury frequency standard based on a single laser-cooled $^{199}\text{Hg}^+$ ion. This device makes use of an energy transition that provides a frequency of 1.06×10^{15} Hz with a line width of 6.7 Hz (Rafac *et al.*, 2000) from a single ion stored in a cryogenic Paul ion trap. This transition is in the ultraviolet (wavelength = 282 nm) and results in a Q value of 1.6×10^{14} . By comparing this standard

to an optical stand based on neutral calcium atoms operating with a transition frequency of 456 THz (4.56×10^{14}) it was determined that the mercury device was capable of a 1-second stability of 1×10^{-15} (Diddams *et al.*, 2001). Other ions have also been investigated for such applications, including ytterbium, strontium and aluminum.

11.3

Optical Neutral Atom Clocks

Neutral atoms can also be used in optical frequency standards. These clocks make use of optical standard technology and combine the advantages of using a trapped single ion with using a large number of free-falling neutral atoms. The performance of a single-ion clock is compromised because the signal-to-noise ratio is low, which is because it only involves one ion. Using a cloud of neutral atoms in an optical clock is a possible means to get around the problem. Calcium has been used in this capacity because of a convenient wavelength at 657 nm and its insensitivity to external fields (Oates, Bondu and Hollberg, 1999). Millions of calcium atoms are cooled in a magneto-optic trap, released and then probed using laser light at 657 nm. A femtosecond mode-locked laser is used to provide a femtosecond comb, which, in turn, is used to provide the frequency of the clock transition.

Neutral atoms are also used in optical clocks by confining them in an optical lattice. To take advantage of some of the narrow line widths that are possible optically, it is necessary to increase the interrogation time of the atoms. Katori *et al.* (2003) proposed using an optical lattice to confine the atoms so that the interrogation times could be increased. An optical lattice makes use of standing light waves to create a potential surface that is made up of a series of ‘valleys’ in which the atoms are confined (Figure 11.3). The crystal-like structure in which the atoms are confined makes it possible to increase the interrogation time of the atoms. The interfering light beams that create the lattice must operate at a wavelength such that the light shifts that are exerted on the ground and upper states of the clock transition are exactly equal. This ‘magic frequency’ prevents the optical field that creates the lattice from perturbing the clock transition frequency.

Neutral atoms are first slowed and cooled and then loaded into the optical lattice. Atoms of strontium (Ludlow *et al.*, 2008) and ytterbium (Barber *et al.*, 2008) have been used in these standards.

11.4

Quantum Logic Clock

Quantum logic spectroscopy has opened up another approach to a frequency standard. This technique uses a single ion of one element along with a single ion

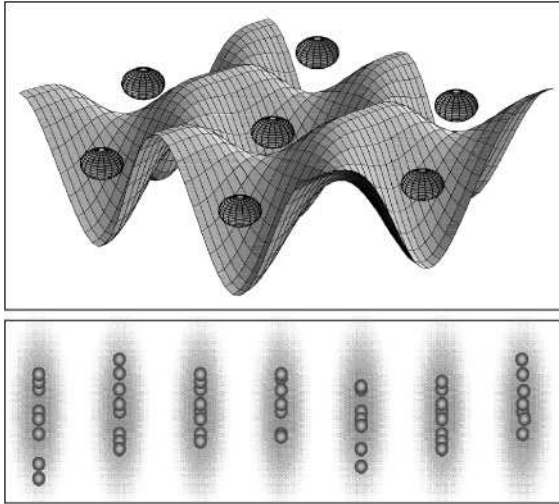


Figure 11.3 An optical lattice in three dimensions (top) and two dimensions (bottom).

of an auxiliary element to handle the requirements for laser cooling and state detection. Such a standard has been developed using an aluminum $^{27}\text{Al}^+$ ion and a beryllium $^9\text{Be}^+$ ion (Rosenband *et al.*, 2007). Both are loaded into a linear Paul trap where they can be expected to exist as an ion pair for several hours. A chemical interaction with background gas eventually removes one of the ions from the trap. The pair acts as a crystal within the trap. Lasers operating at a wavelength of 313 nm act on the beryllium to cool the ion pair. A second laser is used to get the aluminum ion to the proper state where the light from yet another laser at a wavelength of 267.4 nm is able to excite the clock transition. When the aluminum ion is in the excited state, the beryllium ion is constrained, by Coulomb interaction, to be in the level to which it had been previously pumped. The fluorescence of photons from this level can be monitored to provide the feedback signal to the excitation laser. The standard is operated by a series of pulses of this process. In each pulse seven photons are expected to be seen if the aluminum ion is in the proper state, and only one is expected to be seen if the clock transition was not successful.

The quantum logic standard receives this name because it uses techniques used in quantum computers that are based on quantum mechanics. There are two different logical states of the aluminum ion. The actual state of the ion is communicated to the beryllium ion, which then provides easily detected signals depending on its state. The optical frequency of the excitation laser can be translated to rf frequencies using a femtosecond comb as outlined previously. Aluminum is used in this process because it provides a stable frequency. Because it is difficult to detect, however, the beryllium ion is introduced.

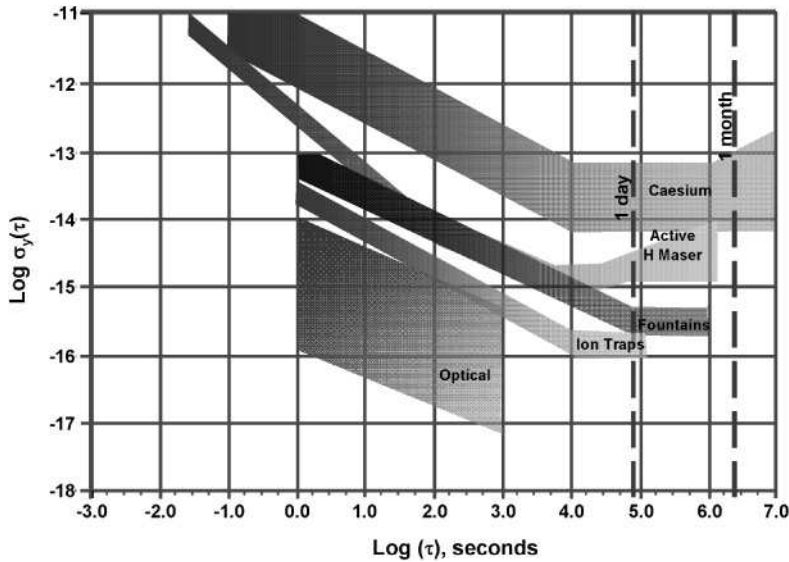


Figure 11.4 Stabilities of microwave and optical timekeeping standards.

11.5 Characterizing Optical Standards

Optical standards offer great promise for precise timekeeping (Gill, 2005; Rosenband *et al.*, 2007). Their expected contributions are shown in Figure 11.4.

References

- Barber, Z.W., Stalnaker, J.E., Lemke, N.D., Poli, N., Oates, C.W., Fortier, T.M., Diddams, S.A., Hollberg, L., Hoyt, C.W., Taichenachev, A.V. and Yudin, V.I. (2008) Optical lattice induced light shifts in an Yb atomic clock. *Phys. Rev. Lett.*, **100**, 103002.
- Dehmelt, H.G. (1982) Mono-ion oscillator as potential ultimate laser frequency standard. *IEEE Trans. Instrum. Meas.*, **31**, 83–7.
- Diddams, S.A., Hollberg, L., Ma, L.-S. and Robertson, L. (2002) Femtosecond-laser-based optical clockwork with instability less than or equal to 6×10^{-6} in 1 s. *Opt. Lett.*, **27**, 58–60.
- Diddams, S.A., Udem, T., Bergquist, J.C., Curtis, E.A., Drullinger, R.E., Hollberg, L., Itano, W.M., Lee, W.D., Oates, C.W., Vogel, K.R. and Wineland, D.J. (2001) An optical clock based on a single trapped $^{199}\text{Hg}^+$ ion. *Science*, **293**, 825–8.
- Gill, P. (2005) Optical frequency standards. *Metrologia*, **42**, S125–37.
- Holzwarth, R., Udem, T., Hänsch, T.W., Knight, J.C., Wadsworth, W.J. and St. Russell, P. (2000) Optical frequency synthesizer for precision spectroscopy. *Phys. Rev. Lett.*, **85**, 2264–7.
- Katori, H., Takamoto, M., Pal'chikov, V.G. and Ovsiannikov, V.D. (2003) Ultrastable optical clock with neutral atoms in an engineered light shift trap. *Phys. Rev. Lett.*, **91**, 173005–8.
- Ludlow, A.D., Zelevinsky, T., Campbell, G.K., Blatt, S., Boyd, M.M., de Miranda, M.H.G., Martin, M.J., Thomsen, J.W., Foreman,

- S.M., Ye, J., Fortier, T.M., Stalnaker, J.E., Diddams, S.A., Le Coq, Y., Barber, Z.W., Poli, N., Lemke, N.D., Beck, K.M. and Oates, C.W. (2008) Sr lattice clock at 1×10^{-16} fractional uncertainty by remote optical evaluation with a Ca clock. *Science*, **319**, 1805–8.
- Ma, L.-S., Bi, Z., Bartels, A., Robertsson, L., Zucco, M., Windeler, R.S., Wilpers, G., Oates, C., Hollberg, L. and Diddams, S.A. (2004) Optical frequency synthesis and comparison with uncertainty at the 10–19 level. *Science*, **303**, 1843–5.
- Oates, C.W., Bondu, F. and Hollberg, L. (1999) A diode-laser optical frequency reference based on laser-cooled Ca atoms. *Eur. Phys. J. D*, **7**, 449–59.
- Rafac, R.J., Young, B.C., Beall, J.A., Itano, W.M., Wineland, D.J. and Bergquist, J.C. (2000) Sub-dekahertz ultraviolet spectroscopy of $^{199}\text{Hg}^+$. *Phys. Rev. Lett.*, **85**, 2462–5.
- Rosenband, T., Schmidt, P.O., Hume, D.B., Itano, W.M., Fortier, T.M., Stalnaker, J.E., Kim, K., Diddams, S.A., Koelemeij, J.C.J., Bergquist, J.C. and Wineland, D.J. (2007) Observation of the $^1\text{S}_0 \rightarrow ^3\text{P}_0$ clock transition in $^{27}\text{Al}^+$. *Phys. Rev. Lett.*, **98**, 220801.
- Udem, T., Reichert, J., Holzwarth, R. and Hänsch, T.W. (1999) Accurate measurement of large optical frequency differences with a mode-locked laser. *Opt. Lett.*, **24**, 881–3.

12

Definition and Role of a Second

12.1

The Historical Second

Early time keepers had little need for time divisions finer than an hour or, at most, a simple fraction of an hour. In geometry, the circle was divided into 360 degrees in the last centuries BC by Babylonian astronomers, but they had developed the sexagesimal system much earlier for nonastronomical use (Neugebauer, 1957). The origin of the concept of 360 degrees in a circle is not clear. Some suspect that it is related to the length of the year in days. Another possibility is the fact that a hexagon composed of six equilateral triangles can be inscribed within a circle. Then, if each of the angles of the triangles could be described by 60 degrees, following the sexagesimal system, the circle would have 360 degrees. The first known Hellenistic geometer to make use of these divisions was Hupsikles around 150 BC (Irby-Massie and Keyser, 2002). Claudius Ptolemy (about 100–175 AD) followed the custom and used finer subdivisions of the degree in his work *Mathematike Syntaxis*, now known as the *Almagest*, the title being a Latin form of the Arabic translation *al-kitabu-l-mijist*. The first Latin translation of Ptolemy's treatise did not become available until the twelfth century, when the work by Gerard of Cremona was published (McCluskey, 1998). In it we find the Latin translation of the subdivisions of the degree used by Ptolemy as *partes minutae primae*, or first minutes, which became known simply as 'minutes' and the subsequent further subdivision of that unit, *partes minutae secundae*, or 'second minutes,' which became known as 'seconds.' The use of these angular units was restricted mainly to theoretical and astronomical usage.

In everyday common time-keeping usage the hour was essentially divided into halves, thirds, quarters, or sometimes twelfths until the end of the sixteenth century, but not into 60 minutes. In dealing with time, Claudius Ptolemy had divided the day either into four parts, each containing six hours, or into 360 'chronoi.' Consequently each hour was made up of 15 chronoi. Venerable Bede divided the hour into 4 'puncti,' or 10 'minuta,' or 15 'partes,' or 40 'momenta.' He considered the punctum to be the smallest unit that could be measured with a sundial and the momentum as the smallest conceivable unit (Dohrn-van

Rossum, 1996). In the medieval world, the word 'minutum' was also used in various writings to indicate 1/15 of an hour, 1/10 of an hour, and 1/60 of a day. The word 'ostentum' was used to indicate 1/60 of an hour (Holford-Strevens, 2005). Other units of time used were saeculum (century), lustrum (five years), annus (year), mensis (month), hebdomada (week), dies (day), hora (hour), quadrans (quarter of an hour), minutum (minute), momentum (the length of time needed to discern that time has moved on), ostentum (the time needed to take something in visually), ictus oculi (twinkling of an eye) which was equated with the atomus (the unit of time that could not be divided further) (Fuhrman, 1986).

In the later Middle Ages, however, the concept of minutae primae, secundae and even tertiae had arisen in the face of this lack of standardization of time units. Such terminology had been used in referring to measures of arc, but was being applied to time units as well. Although the concept of an hour with 60 minutes of 60 seconds each was understood by the mid-fourteenth century, the concept still did not find widespread common usage because clocks were not reliably capable of providing the finer time units. Minutes are mentioned in the fourteenth century, and clocks that indicated minutes may have existed by the end of the fifteenth century, but there is no reliable evidence to support that supposition. Seconds were not considered in common practical timekeeping for at least another century. In the late seventeenth century, following the development of pendulum clocks and the anchor escapement (Chapter 9) clocks with minute hands began to appear (Milham, 1923), and clocks with second hands followed in the eighteenth century.

In the area of precise timekeeping for scientific usage, Tycho Brahe mentions the use of minutes in his journals only twice between 1563 and 1570. In 1577 he refers to new clocks showing minutes and in 1581 he refers to seconds. In 1587 he complained about the fact that his clocks could not be made to agree to better than four seconds (Landes, 1983). In 1579, a Swiss clockmaker, Jost Bürgi, at the court of William of Hesse is reported to have developed a clock that marked seconds as well as minutes. Further, this clock is said to have been precise at the level of a minute per day (Landes, 1983).

In the early seventeenth century navigation at sea over large distances provided some incentive for more precise timekeeping. At that time the determination of latitude was relatively well known, but accurate measurement of longitude remained a problem. Typically mariners would sail toward the parallel of latitude of their destination and sail along that until their port was in sight. Their progress in longitude was measured by dead reckoning, which was aided by estimates of their speed in the water. To do that they made use of a 'log' attached to a knotted rope. The log was thrown overboard and the knots that passed through the mariner's hands during a specified time duration were counted in order to determine speed. That time interval might be measured by the length of time required to recite a prayer or a verse from a song, but in the Middle Ages the sand glass began to be used to measure those intervals of time (Landes, 1983). Sandglasses came

into being at about the same time as wheeled clocks in the late thirteenth century and were becoming familiar instruments to measure short time intervals in the fourteenth century (Dohrn-van Rossum, 1996). Although they are called ‘sandglasses,’ other materials were often used including powdered rock or crushed egg shells (Lippincott, 1999).

The definition of the second was assumed to be the sixtieth part of a minute which was the sixtieth part of an hour, and nothing more formal than that was required. Tito Livio Burattini (1617–1681) (Burattini, 1675) perhaps provides the first formal definition of the second as $1/86400$ of a solar day (Leschiutta, 2005). This appears in his work *Misura universale* [Universal measure] (1675), where he suggests a standard length unit equivalent to the length of a pendulum with a period of one second. This definition of a second appears to have been sufficient for all practical purposes into the twentieth century. No formal definition of the second beyond the appropriate fraction of a day appears to have been needed. The organization responsible for world metrology surprisingly never formally defined the second of mean solar time (Audoin and Guinot, 2001).

12.2

The Ephemeris Second

A definition of a second based on the Earth’s variable rotation became impractical for precise time-keeping applications in the twentieth century. Chapter 6 outlines the development of the concept of Ephemeris Time (ET) that led to the first modern formal definition of the second, which was proposed to address this problem. The tenth meeting of the Conférence Générale des Poids et Mesures, (CGPM) in 1954, following the earlier recommendation of the Comité International des Poids et Mesures (CIPM), proposed the following formal definition of the second: ‘The second is the fraction $1/31\,556\,925.975$ of the length of the tropical year for 1900.0.’ (Trans. Int. Astron. Union, 1957).

As seen in Chapter 6, this fraction was based on Newcomb’s formula for the geometric mean longitude of the Sun for the epoch of January 0, 1900, 12h UT (Newcomb, 1895) given by

$$L = 279^\circ 41' 46.04'' + 129\,602\,768.13'' T + 1.089'' T^2, \quad (12.1)$$

where T is the time reckoned in Julian centuries of 36525 days.

In 1956 the CIPM adopted the slightly more precise value with the words: ‘La seconde est la fraction $1/31\,556\,925.9747$ de l’année tropique pour 1900 janvier 0 a 12 heures de temps des ephemerides.’ It also created the Comité Consultatif pour la Définition de la Seconde (CCDS) to coordinate future work in the area (Procès Verbaux des Séances, 1957). The ephemeris second, as a fraction of the tropical year, was formally adopted by the 11th CGPM in 1960.

12.3

The SI Second

Events leading to the definition of the second of the *Système International* (SI) began with the development of the microwave cesium frequency standard (Chapter 10), which took place concurrently with the discussion and adoption of the Ephemeris second. Following the establishment of the possible viability of the cesium standard as a clock, it became necessary to calibrate the frequency of the atomic transition and thus establish its time scale unit in relation to the prevailing definition of the second. Essen and Parry (1957) report that values of the frequency of cesium were reported by Sherwood *et al.* (1952) and in 1956. The 1952 value was $9\,192\,632\,000 \pm 2000$ Hz, and the 1956 values were $9\,192\,631\,970 \pm 90$ Hz, $9\,192\,631\,800 \pm 50$ Hz, and $9\,192\,631\,880 \pm 30$ Hz. These were consistent with the Essen and Parry (1955) value of $9\,192\,631\,830 \pm 10$ Hz and the Essen and Parry (1957) value of $9\,192\,631\,845 \pm 2$ Hz, both of which were determined with respect to the second of Universal Time (UT2) determined using astronomical observations of the Royal Greenwich Observatory.

The actual definition used for the SI second, however, was the result of the collaboration of the UK physicists L. Essen and J.V. L. Parry with the U.S. astronomers Wm. Markowitz and R. G. Hall (see Section 10.4.1.1). From this collaboration, Markowitz *et al.* (1958) arrived at a final value of $9\,192\,631\,770 \pm 20$ Hz. The second in this determination was the second of Ephemeris Time that was determined from astronomical observations of the Moon with respect to a star background. Leschiutta (2005) states that

As a personal remark, taking into account the capabilities of the timing emissions at the moment, of the frequency standards available, of the inevitable scatter of the moon camera, and some other factors, not least the widespread use and abuse in ‘touching’ the piezo-oscillator, it is almost impossible to explain the accuracy of the Markowitz determination. Similar events, i.e. results surpassing the capabilities of the moment, are not uncommon in the history of science that sometimes is prone to accepting the intervention of a serendipity principle. The other possible explanation calls for a first class understanding of physics, coupled with scientific integrity. ...

The question about the mental and experimental paths taken by Markowitz that led him to write $(9\,192\,631\,770 \pm 10)$ Hz remains open, despite:

- using UT1 data with clocks corrected nearly every day, with a timing accuracy at the millisecond level at best, giving an accuracy in frequency of 1×10^{-8} ;
- taking data from the double-rate moon camera, having a resolution of around 0.5 s on ET and for UT from a PZT affected by a resolution of 10 ms;

- the need for a long chain of measurements and commands, composed of many laboratories belonging to different authorities, to be kept ‘synchronized’.

The Markowitz *et al.* (1958) publication does not provide extensive details of the process by which they arrived at this number that was destined to become the definition of the SI second. The frequency of the cesium hyperfine transition that defines the second is actually discussed in two papers by the collaborators (Essen *et al.*, 1958; Markowitz *et al.* 1958). The Essen *et al.* (1958) paper describes the calibration of the cesium transition in terms of time determined from the Earth’s rotation. The second paper describes the correction to this frequency to put it in terms of the second of Ephemeris Time. Leschiutta (2005) attempts to fill in some of the details based on subsequent discussions.

Two piezoelectric quartz clocks were involved in the process. The first was the device used to provide the frequency required to produce the hyperfine transition in the cesium atoms of Essen’s standard in the United Kingdom. So, at the National Physical Laboratory (NPL) in the United Kingdom, a quartz clock was adjusted in rate to the frequency provided by the cesium standard. The second quartz oscillator needed was one at the U.S. Naval Observatory (USNO) in Washington that was regulated by astronomical observations of the Earth’s rotation.

In the 1950s the instruments used to make the astronomical observations for timekeeping at USNO were the Photographic Zenith Tubes (PZT) at Washington and at Richmond (outside Miami), Florida. The PZT was a telescope having a lens with an eight-inch diameter and a very long focal length. It was constrained to look only at the zenith and designed to make use of a mercury pool to define the direction of the vertical. The pool was inserted midway in the light path and reflected the light to a photographic plate immediately underneath the lens which recorded the stellar images. The telescope was capable of determining the difference in time between the astronomically determined time (Universal Time) and the clock time daily with an accuracy of a few milliseconds (see Figure 12.1).

The USNO clock, then, was regulated to provide a realization of UT2, which is the astronomically determined time corrected for the known seasonal variation in the Earth’s rotation (see Chapter 14). Its frequency can be denoted as ν_{UT2} and the clock reading after a time interval τ can be written $h_{\text{UT2}} = \nu_{\text{UT2}}\tau$. Similarly, after the same interval, the NPL clock, whose frequency ν_{CS} was regulated to provide time based on the frequency of the nominal cesium frequency, would read $h_{\text{CS}} = \nu_{\text{CS}}\tau$. To calibrate the cesium frequency, then, in terms of UT2, it was necessary to determine ν_{UT2} by comparing the frequencies of the two quartz clocks at opposite ends of the Atlantic Ocean. Each facility monitored the same timing signals broadcast by radio stations in order to determine the difference in time between the two clocks. The stations used for this were WWV in Washington, DC, United States, and GBR in Rugby, United Kingdom, 120km away from NPL. In the Essen *et al.* (1958) analysis, smoothed monthly differences between the two clocks were used to determine the relative frequency difference. The smoothing of the astronomical

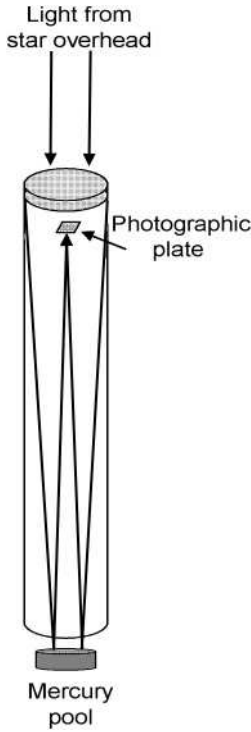


Figure 12.1 Photographic Zenith Tube.

observations also eliminated the effects of known tidal variations with periods of 27.6 and 13.6 days in the Earth's rotation. The radio signals were monitored near the beginning of each month from June, 1955, through January, 1958.

Mathematically, if each institution observes the same radio time signal at the same time at the beginning and end of the month, it is possible to determine the differences h_{UT2} and h_{CS} . Of course, corrections to the actual observations of the time differences must be made to allow for the times of transmission of the radio signals to the respective monitoring locations. Then after making the systematic corrections, it is possible to write

$$h_{\text{CS}} - h_{\text{UT2}} = \Delta h = (v_{\text{CS}} - v_{\text{UT2}}) \tau, \quad (12.2)$$

$$\text{but } \tau = \frac{h_{\text{UT2}}}{v_{\text{UT2}}}, \quad (12.3)$$

$$\text{so } \Delta h = (v_{\text{CS}} - v_{\text{UT2}}) \frac{h_{\text{UT2}}}{v_{\text{UT2}}}, \quad (12.4)$$

$$\text{and } v_{\text{UT2}} = v_{\text{CS}} - v_{\text{UT2}} \frac{\Delta h}{h_{\text{UT2}}}. \quad (12.5)$$

Following this process, Markowitz *et al.* (1958) reported that they had found the frequency of cesium in terms of the UT2 second to be

$\nu_{\text{UT2}} = 9\,192\,631\,882$ cycles per second of UT2.

Observations made with the Markowitz Moon Camera (see Chapter 6) at Washington were then used to determine the difference between the Ephemeris Time second and the UT2 second. The same process that was used to determine ν_{UT2} could be used to determine ν_{ET} , the frequency of the cesium energy level transition in terms of the ephemeris second. A time series of the differences between Ephemeris Time and UT2 could be used in place of the time series of the differences between a clock reading UT2 and a clock operating with a nominal cesium frequency to estimate

$$\nu_{\text{ET}} = \nu_{\text{UT2}} - \nu_{\text{ET}} \frac{\Delta h}{h_{\text{ET}}}. \quad (12.6)$$

Their result was (Markowitz *et al.*, 1958)

$\nu_{\text{ET}} = 9\,192\,631\,770 \pm 20$ cycles per second of Ephemeris Time.

12.4

Adopting the SI Second

The SI brochure (BIPM, 2006) summarizes the transition to the atomic second defined by Markowitz, Hall, Essen and Parry:

In order to define the unit of time more precisely, the 11th CGPM (1960, **Resolution 9**) adopted a definition given by the International Astronomical Union based on the tropical year 1900. Experimental work, however, had already shown that an atomic standard of time, based on a transition between two energy levels of an atom or a molecule, could be realized and reproduced much more accurately. Considering that a very precise definition of the unit of time is indispensable for science and technology, the 13th CGPM (1967/68, Resolution 1) replaced the definition of the second by the following:

The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium 133 atom.

It follows that the hyperfine splitting in the ground state of the cesium 133 atom is exactly 9 192 631 770 hertz, $\nu(\text{hfs Cs}) = 9\,192\,631\,770 \text{ Hz}$. [The symbol, (hfs Cs), is used to denote the frequency of the hyperfine transition in the ground state of the cesium atom.]

At its 1997 meeting the CIPM affirmed that:

This definition refers to a cesium atom at rest at a temperature of 0 K.

This note was intended to make it clear that the definition of the SI second is based on a cesium atom unperturbed by black body radiation, that is, in an environment whose thermodynamic temperature is 0 K. The frequencies of all primary frequency standards should, therefore, be corrected for the shift due to ambient radiation, as stated at the meeting of the Consultative Committee for Time and Frequency in 1999.

The text of the 13th CGPM Resolution that defined the SI second is given in the brochure:

‘The 13th Conférence Générale des Poids et Mesures (CGPM),

Considering

- that the definition of the second adopted by the Comité International des Poids et Mesures (CIPM) in 1956 (Resolution 1) and ratified by Resolution 9 of the 11th CGPM (1960), later upheld by Resolution 5 of the 12th CGPM (1964), is inadequate for the present needs of metrology,
- that at its meeting of 1964 the CIPM, empowered by Resolution 5 of the 12th CGPM (1964), recommended, in order to fulfill these requirements, a cesium atomic frequency standard for temporary use,
- that this frequency standard has now been sufficiently tested and found sufficiently accurate to provide a definition of the second fulfilling present requirements,
- that the time has now come to replace the definition now in force of the unit of time of the *Système International d’Unités* by an atomic definition based on that standard,

Decides

1. The SI unit of time is the second defined 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 cesium 133 atom;’
2. Resolution 1 adopted by the CIPM at its meeting of 1956 and Resolution 9 of the 11th CGPM are now abrogated.’

In the world of international standards, the definition of the second directly affects the SI base unit of the meter, which is now defined thus:

‘The meter is the length of the path traveled by light in vacuum during a time interval of $1/299\,792\,458$ of a second.’

It also indirectly affects the definition of the SI base unit of the ampere, which is:

‘The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible

circular cross-section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length.'

as well as the candela, which is defined thus:

'The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian.'

Numerous other derived SI units also involve this definition of the second. It should also be noted that the SI abbreviation for the unit of a second is 's'.

In less than 30 years, from 1954 to 1983, we went from no formal definition of the second to the definition of the second being the most precise metrological unit and the basis for other units including the definition of the meter in connection with the speed of light.

References

- Audoin, C. and Guinot, B. (2001) *The Measurement of Time*, Cambridge University Press, Cambridge.
- BIPM (2006) *The International System of Units (SI)*, 8th edn, BIPM.
- Burattini, T.L. (1675) *Misura universale*, Stamperia dei Padri Francescani, Vilna.
- Dohrn-van Rossum, G. (1996) *History of the Hour* (transl. T. Dunlap), University of Chicago Press, Chicago.
- Essen, L. and Parry, J.V.L. (1955) An Atomic Standard of Frequency and Time Interval: A Caesium Resonator, *Nature*, **176**, 280–2.
- Essen, L. and Parry, J.V.L. (1957) The Caesium Resonator as a Standard of Frequency and Time, *Phil. Trans. R. Soc. London, Ser. A, Math. Phys. Sci.*, **250**, 45–69.
- Essen, L., Parry, J.V.L., Markowitz, W. and Hall, R.G. (1958) Variation in the speed of rotation of the earth since June 1955, *Nature*, **181**, 1054.
- Fuhrman, H. (1986) *History of Germany in the High Middle Ages. 1050–1200* (transl. T. Reuter), Cambridge University Press.
- Holford-Strevens, L. (2005) *The History of Time, A Very Short Introduction*, University Press, Oxford, p. 144.
- Irby-Massie, G.L. and Keyser, P.T. (2002) *Greek Science of the Hellenistic Era*, Routledge, London, p. 392.
- Landes, D.S. (1983) *Revolution in Time, Clocks and the Making of the Modern World*, Belknap Press of the Harvard University Press, Cambridge, MA and London.
- Leschiutta, S. (2005) The definition of the 'atomic' second. *Metrologia*, **42**, S10–19.
- Lippincott, K. (1999) *The Story of Time*, Merrell Holberton, London, p. 304.
- McCluskey, S.C. (1998) *Astronomies and Cultures in Early Medieval Europe*, Cambridge University Press, p. 231.
- Markowitz, W., Hall, R.G., Essen, L. and Perry, J.V.L. (1958) Frequency of cesium in terms of ephemeris time. *Phys. Rev. Lett.*, **1**, 105–7.
- Milham, W.I. (1923) *Time & Timekeepers Including the History, Construction, Care, and Accuracy of Clocks and Watches*, Macmillan, New York and London.
- Neugebauer, O. (1957) *The Exact Sciences in Antiquity*, 2nd edn, Brown University Press, Providence.
- Newcomb, S. (1895) *Astronomical Papers Prepared for the Use of the American Ephemeris and Nautical Almanac*, Vol. VI, Part I: *Tables of the Sun*, U.S. Gov. Printing Office, Washington, D.C.
- Procès-Verbaux of the Comité International des Poids et Mesures (1957) *deuxième série*, **25**, 77.

- Sherwood, J.E., Lyons, H., McCracken, R.H. and Kusch, P. (1952) High frequency lines in the hfs spectrum of cesium. *Bull. Am. Phys. Soc.*, **27** (I), 43.
- Trans. Int. Astron. Union (1957) *Proc. 9th General Assembly, Dublin, 1955*, Vol. IX (ed. P.T. Oosterhoff), Cambridge University Press, New York, p. 451.

13

International Atomic Time (TAI)

13.1

Constructing an Atomic Time Scale

The advent of a time scale based on atomic energy level transitions soon followed the development of the new technology of atomic clocks. In his *Time for Reflection*, Louis Essen (2000) described the situation following his success with the first cesium standard:

Our first task was to make every possible test to check to what extent the frequency could be varied by external conditions such as pressure, temperature, strength of the electric and magnetic fields, and so on. This could only be done by establishing a provisional atomic time scale, making use of the stability of our quartz clocks. They were set at intervals of a week by means of the atomic clock operating under standard conditions. These conditions were then varied and the effect measured by the quartz clocks. The test showed that with a very simple control of the conditions the atomic clock was enormously more accurate than astronomical time as well as having the advantages of being far simpler to use and being immediately available. It did not give the time of day, of course, but this is not required accurately. It is the length of a time interval and its inverse, frequency, that is needed ever more accurately for modern developments in navigation, computers and communication.

The earliest atomic time scales were constructed using a single atomic frequency standard to steer a quartz crystal clock. In 1955, shortly after the appearance of the first operational cesium frequency standard at the National Physical Laboratory (NPL) in the United Kingdom (Essen and Parry, 1957), the Royal Greenwich Observatory (RGO) established its time scale, called “Greenwich Atomic” (GA). It was formed using quartz crystal clocks whose frequencies were calibrated at periodic intervals with the NPL cesium frequency standard. On 13 September 1956 the U.S. Naval Observatory began its atomic time scale, called ‘A.1,’ again using a quartz crystal clock calibrated daily with an Atomichron® (see Section 10.4.1.2)

cesium beam standard located at the U.S. Naval Research Laboratory (NRL) in Washington (*Time Service Notice No. 6*, 1959). In neither case were the cesium standards operated continuously (Nelson *et al.*, 2001). The atomic frequency standards were switched on periodically for a short duration in order to calibrate the frequency of the quartz crystal and then switched off again. Other laboratories and institutions followed quickly in constructing atomic time scales, including the U.S. National Bureau of Standards (NBS) in Boulder, Colorado, which began an atomic time scale, NBS-A, on 9 October 1957. The epoch of this time scale as well as that of A.1 were set equal to the astronomically determined time UT2 on 1 January 1958 (Barnes, Andrews and Allan, 1965).

As more standards became available, constructing a time scale to take advantage of several standards became a challenge. In constructing such a time scale some consideration needs to be given to the characteristics of the final product. These considerations include concerns about reliability, frequency stability, frequency accuracy, and accessibility (Guinot and Arias, 2005). The reliability of the scale depends critically on the reliability of the clocks whose data are used to produce the scale. To ensure this reliability it is desirable to have a large number of clocks contributing data, thereby reducing the dependence of the scale on a few contributors. The frequency stability refers to the ability of the scale to maintain a scale interval with a constant relationship with respect to the conventionally adopted international standard second. To be stable the scale need not necessarily be accurate, but it must have a ratio of its scale interval to the standard second that is as constant as possible. Allan Variance is used to measure stability quantitatively. Frequency accuracy, on the other hand, requires that the scale interval be as close as possible to the standard second. Accessibility is provided by the clocks whose data are used in the formation of the time scale. In the process of computing the time scale, the arithmetic differences between the contributed clock time and the time scale can be provided to the contributors, who, in turn, can provide that to their client users of precise time.

Thought also must be given to the desired period over which the frequency of the resultant time scale is to be stabilized, as well as the period of time between the contributed comparisons of the clock data. Still other considerations include how often the time scale is to be computed and the time between the last contributed data and the time when the product is made available (Audoin and Guinot, 2001).

The USNO expanded its A.1 scale by making use of cesium standards at a number of cooperating institutions. By 1961 the A.1 scale was constructed using data from cesium frequency standards located at USNO, NRL, NBS, NPL, Harvard University, National Research Council of Canada in Ottawa, Centre National d'Études des Télécommunications (Bagneux, France), Observatoire de Neuchâtel (Switzerland), and the USNO Time Service Sub-station located at Richmond Heights, Florida, just south of Miami. The frequency comparisons were done by monitoring the phase of the signals from the very-low-frequency (VLF) radio stations broadcasting with frequencies between 3 and 30 KHz. (Markowitz, 1962; Trans. Int. Astron. Union, 1962).

13.2 History of TAI

In July, 1955, the Bureau International de l'Heure (BIH) began an atomic time scale that has been continuous ever since. From 1955 through 1969 the BIH used VLF phase comparison data from distant atomic clocks and comparisons of local cesium standards to provide a mean time scale, called T_m or AM. All of these data were referenced to a quartz or rubidium clock at the Observatoire de Paris where the BIH was located. Just as with A.1, this time scale was set equal to UT2 at 0h, 1 January 1958. Beginning in 1960, the BIH routinely published the difference between AM and astronomically observed UT2 in periodic bulletins.

In 1963 the BIH decided that, instead of giving all contributing clocks equal weight in the mean, it would be better to use as input only the three standards located at the metrology laboratories at the U.S. National Bureau of Standards, the Swiss Laboratoire Suisse de Recherches Horlogères, and the National Physical Laboratory in the United Kingdom. The time scale then became known as A3 in recognition of the contributions of the three institutions. In 1966 the scale was expanded to make use of contributions from other laboratories, but the name A3 was still used (Guinot and Arias, 2005; Nelson *et al.*, 2001).

Technological advances in time transfer, particularly the development of LORAN-C and time transfer via television signals, along with the advent of portable cesium clocks to improve calibrations, allowed the BIH to improve its time scale significantly. Consequently, it changed its procedure to average contributing independent local time scales, instead of the contributing clock data. The time scale created using this process began in 1969 and was designated TA(BIH). It was constrained to be continuous with A3. The initial time scales, designated TA(k) with k representing a laboratory identifier, that were included in TA(BIH) were those of the Physikalische Technische Bundesanstalt (PTB) of Germany, the Commission Nationale de l'Heure of France (F), and the U.S. Naval Observatory (USNO). During 1969 the Royal Greenwich Observatory (RGO), the National Research Laboratory of Canada (NRC), the National Bureau of Standards (NBS), and the Observatoire de Neuchâtel (ON) in Switzerland were added to the list of contributors (Guinot and Arias, 2005).

The timekeeping community was able to use the BIH time scale by means of the published values of TA(BIH)–TA(k) that were available in monthly BIH circulars. Institution k could then determine the time varying relation of its time scale TA(k) with respect to the standard TA(BIH) and make any appropriate changes in its scale to comply more closely with that of the BIH. The TA(BIH)–TA(k) data were available with a delay of one to two months and with uncertainties of 10 μ s for those stations using VLF frequency comparisons and 1 μ s for those using LORAN-C (Guinot and Arias, 2005).

Despite initial reluctance to accept time determined by means other than from astronomical observations of the Earth's rotation, atomic time began to be accepted gradually. Essen describes the gradual acceptance in the following remarks:

A sub-committee of the International Committee of Weights and Measures was set up to discuss atomic time and it is interesting to follow its gradual and reluctant acceptance by astronomers. The meeting in 1957 refused to accept the term atomic clock insisting that it was simply a frequency standard for the second: the second meeting in 1961 accepted that it was a standard of time interval but continued to stall by recommending that further work should be done: the third meeting in 1963 recommended the adoption of an atomic unit of time the value being that obtained at the NPL. No formal steps were taken to implement this recommendation and the International Scientific Radio Union, in which I was the chairman of the relevant section, had to stress the urgency of putting the resolution into effect. It was formally adopted as the unit of time in 1968 with only one abstention, the representative of the Greenwich Observatory, I regret to say.

The International Astronomical Union (IAU) recommended the unification of time through an atomic time scale in 1967. This was followed by similar recommendations of the International Union of Radio Sciences in 1969 and the International Radio Consultative Committee, now the International Telecommunication Union, in 1970 (Guinot and Arias, 2005). In June, 1970, the Comité Consultatif pour la Définition de la Seconde (CCDS), discussed the need for an atomic scale and that the Bureau International des Poids et Mesures (BIPM) should deal not only with the definition of the time scale interval (the duration of the second), but also with time scales. Consequently they submitted two resolutions to the Comité International des Poids et Mesures (CIPM) on the subject of time scales. The first one pointed out the need for an international atomic time scale to coordinate time signals, to serve as a uniform time reference especially for the dynamics of natural or artificial celestial objects, and to compare frequency standards operating in different places or times. The second suggested a definition of international atomic time (Terrien, 1970). These were then presented to the Conférence Générale des Poids et Mesures (CGPM), which met in October, 1971 and passed the following resolutions (*Comptes Rendus de la 14^e CGPM* (1971), 1972, 78):

‘The 14th Conférence Générale des Poids et Mesures (CGPM),

Considering

- that the second, unit of time of the *Système International d’Unités*, has since 1967 been defined in terms of a natural atomic frequency, and no longer in terms of the time scales provided by astronomical motions,
- that the need for an International Atomic Time (TAI) scale is a consequence of the atomic definition of the second,
- that several international organizations have ensured and are still successfully ensuring the establishment of the time scales based on astronomical motions, particularly thanks to the permanent services of the Bureau International de l’Heure (BIH),

- that the BIH has started to establish an atomic time scale of recognized quality and proven usefulness,
- that the atomic frequency standards for realizing the second have been considered and must continue to be considered by the Comité International des Poids et Mesures (CIPM) helped by a Consultative Committee, and that the unit interval of the International Atomic Time scale must be the second realized according to its atomic definition,
- that all the competent international scientific organizations and the national laboratories active in this field have expressed the wish that the CIPM and the CGPM should give a definition of International Atomic Time, and should contribute to the establishment of the International Atomic Time scale,
- that the usefulness of International Atomic Time entails close coordination with the time scales based on astronomical motions,

Requests the CIPM

- to give a definition of International Atomic Time,
- to take the necessary steps, in agreement with the international organizations concerned, to ensure that available scientific competence and existing facilities are used in the best possible way to realize the International Atomic Time scale and to satisfy the requirements of users of International Atomic Time.'

'La 14^e Conférence générale des poids et mesures,

Considérant

- qu'une échelle de Temps atomique international doit être mise à la disposition des utilisateurs,
- que le Bureau international de l'heure a prouvé qu'il est capable d'assurer ce service,

rend hommage au Bureau international de l'heure pour l'œuvre qu'il a déjà accomplie,

demande aux institutions nationales et internationales de bien vouloir continuer, et si possible augmenter, l'aide qu'elles donnent au Bureau international de l'heure, pour le bien de la communauté scientifique et technique internationale,

autorise le Comité international des poids et mesures à conclure avec le Bureau international de l'heure les arrangements nécessaires pour la réalisation de l'échelle de Temps atomique international à définir par le Comité international.'

They also endorsed the definition of International Atomic Time as: 'The International Atomic Time is the time reference coordinate established by the Bureau International de l'Heure on the basis of the results given by the atomic clocks working in various establishments in accordance with the definition of the second,

the time unit of the International System of Units' (Terrien, 1971). The abbreviation 'TAI' first appears in a recommendation of the CGPM the following year (*Comptes Rendus de la 15^e CGPM* (1975), 1976, 104; Terrien, 1975).

'The 15th Conférence Générale des Poids et Mesures,

having examined the agreement between the Bureau International de l'Heure and the Bureau International des Poids et Mesures designed to meet the requirements of the users of the International Atomic Time (TAI),

notes with pleasure that TAI is made available to users under satisfactory conditions,

renews its request to the national and international organizations to continue, and if possible to increase, the help they provide to the Bureau International de l'Heure,

and asks the Comité International des Poids et Mesures to maintain their relationship with the Bureau International de l'Heure and its Directing Board with a view to improving the accuracy and the continuity of TAI.'

In 1980 the Comité Consultatif pour la Définition de la Seconde (CCDS) discussed concerns related to TAI that resulted in formal recommendations. The first dealt with how TAI was to be considered in relativistic terms (Giacomo, 1981):

'The Comité Consultatif pour la Définition de la Seconde,

Considering

- that the 14th Conférence Générale has decided to establish an international reference time scale, TAI,
- that the CIPM at its 59th Session has defined TAI accordingly,
- that the BIH is charged with the determination of TAI, which is implemented according to the directives of CCDS ('Mise en pratique du Temps Atomique International', CCDS, 5th Session, 1970, p. S22),
- that the use of TAI requires the application of transformations, commonly referred to as relativistic corrections, for measuring time differences between remote clocks,
- that these corrections require the adoption of a clearly-defined model,

States

- that TAI is a coordinate time scale defined in a geocentric reference frame with the SI second as realized on the rotating geoid as the scale unit,
- that accordingly it can be extended at the present state of the art with sufficient accuracy to any fixed or moving point in the vicinity of the geoid by applying

first-order General Relativity corrections, that is, corrections for gravitational potential and velocity differences and for the rotation of the Earth.’

An explanatory note went on to provide details of the mathematical relativistic time transfer procedures to be used in the formation of TAI. The second recommendation thanked the BIH for its work in providing TAI (Giacomo, 1981):

‘The Comité Consultatif pour la Définition de la Seconde,

Considering

- that the BIH with the support of the CIPM determines the international reference time scale TAI,
- that TAI meets present requirements of the users in applications requiring the highest precision,
- that the BIH continues to study ways of improving TAI,
- that the quality of this work and that of the published data meet the highest scientific standards and are at the limit of present day technological capabilities,
- that it is expected that additional applications of precision time technology will be forthcoming in the future,

Wishes to express its recognition and great appreciation for the excellent work of the BIH and encourages the Director of the BIH to continue and to follow the same guidelines which have so far proved remarkably successful.’

In addition they made the following recommendations to the CIPM which were endorsed at its 69th Session in 1980 (Giacomo, 1981).

Recommendation S 1 (1980)

Algorithms for time-scale computation

The Comité Consultatif pour la Définition de la Seconde,

Considering

- that TAI should be as stable and as accurate as possible,
- that the many clocks and frequency standards available have various degrees of stability and accuracy,
- that the uncertainties in the current time comparisons can limit the quality of the computed time scales,
- that only a few primary standards are available to ensure the long-term stability of TAI and its conformity with the definition of the SI second, and

- that the algorithm employed can significantly affect the quality of the resulting time scale,

Recommends

that the development of time scale algorithms to ensure optimum use of the available data be actively pursued.

Recommendation S 2 (1980)

Primary frequency standards

The Comité Consultatif pour la Définition de la Seconde,

Taking note of

- the recommendation S 4 (1974) of the CCDS concerning primary cesium frequency standards, and
- the recommendation A.2 of URSI, August 1978, concerning primary cesium frequency standards, and

Considering

- that the needs are growing for a time scale, TAI, that is stable and accurate in rate,
- that the primary cesium frequency standards are the means to ensure accuracy of rate of TAI,
- that there are many laboratories contributing to the formation of TAI although there are still only three contributing to the accuracy of rate of TAI,
- that primary cesium standards operated as clocks can provide very stable and accurate time scales,
- that frequency standards other than cesium beam standards show significant promise,
- that there is currently insufficient research under way in the area of primary frequency standards,

Recommends

- that more laboratories pursue the development and operation of primary cesium frequency standards, both of conventional and new designs, and
- that laboratories undertake research and development on new frequency standards.

In 1988 the responsibility for the formation of TAI was passed from the BIH to the Bureau International des Poids et Mesures (BIPM) following the recommendation:.. (*Comptes Rendus de la 18^e CGPM* (1987), 1988, 98; Giacomo, 1988).

‘The 18th Conférence Générale des Poids et Mesures

considering the importance of measurements of time and in particular of the International Atomic Time scale, which has already been the subject of Resolution 2 of the 14th Conférence Générale des Poids et Mesures and of Resolutions 4 and 5 of the 15th Conférence Générale,

having taken note of the resolutions adopted by the international Unions concerned—International Astronomical Union, International Union of Geodesy and Geophysics and International Union of Radio Science,

pays tribute to the Bureau International de l’Heure and to its host organization, the Paris Observatory, for creating International Atomic Time and for the quality of the work carried out in order to establish it and diffuse it,

approves the decisions of the Comité International which resulted in assumption by the Bureau International des Poids et Mesures of responsibility for establishing and diffusing International Atomic Time and

recommends the national institutions concerned to pursue with the Bureau International des Poids et Mesures their collaboration for establishing and improving International Atomic Time.’

Although TAI has been acknowledged as an atomic time scale, it has never been disseminated directly nor has it been recognized as the international standard for timekeeping. That distinction is retained for Coordinated Universal Time (UTC) (Chapter 14). Since 1972 TAI and UTC differ by an integral number of seconds, but UTC is the time scale recognized as the international standard.

13.3

Formation of TAI

TAI is a time scale that incorporates the Système International (SI) second to provide users with reliability, frequency stability, frequency accuracy, and accessibility (Guinot and Arias, 2005). The operational concept is to make use of time comparisons of a large number of clocks that are contributed by cooperating institutions from around the world. The objective is to provide a time scale that is more stable than any of the clocks whose data are used in the process. The clock data are combined in a two-step process. The first step is to combine the data to create a preliminary time scale called Échelle Atomique Libre (EAL)

translated into English as ‘Free Atomic Time Scale.’ In a separate step TAI is completed by steering the frequency of the EAL using data from primary frequency standards.

13.3.1

EAL

To provide EAL the BIH developed the algorithm called ‘ALGOS’ to be used to analyze the contributed clock data. Rather than use contributed time scales in the formation of TAI, the BIH decided to use the comparison data from individual contributing clocks. This provided access to more clocks, as some smaller laboratories do not necessarily form a TA(k) time scale from their clocks. Nevertheless the clock data that they provide could be quite useful in the formation of TAI. This also means that the final product, TAI, would be more accessible through a greater number of contributing clocks. Using individual clocks also avoided problems in trying to establish potential systematic adjustments to the TA(k) time scales that institutions might employ to meet their own individual requirements (Guinot and Arias, 2005).

Since its first operational use in 1973 ALGOS (BIH, 1974 and 1975) has been modified a number of times to take advantage of improvements in clocks, frequency standards, and time comparison methods. The principles, however, have remained basically the same since its origin (Guinot and Arias, 2005).

If we assume that at time t , $h_i(t)$ is the reading of clock H_i , then the difference between that reading and EAL(t) can be given by

$$x_i(t) = \text{EAL}(t) - h_i(t), \quad (13.1)$$

and, therefore, that the basic data reported to the BIPM in the form of time differences between clocks at time t are given by:

$$x_{ij}(t) = h_j(t) - h_i(t). \quad (13.2)$$

These quantities can be measured either within a given institution or between laboratories using appropriate techniques for precise time transfer. Currently these comparisons are made using the available time transfer techniques that include Two-Way Satellite Time and Frequency Transfer (TWSFT) and methods that make use of the reception of signals from the Global Positioning System (GPS). Other systems could be used in the future if they can provide the time comparison accuracy. Table 13.1 shows the institutions contributing to the formation of EAL in mid 2008. Figure 13.1 shows the network of laboratory comparison as of mid-2008.

The explanatory note that accompanies the CCDS recommendation of 1980 dealing with the relativistic aspects of TAI (Giacomo, 1981) outlined the basic process of allowing for relativistic effects in the time comparisons. If time com-

Table 13.1 Acronyms (k) and institutions that are possible contributors to the formation of EAL (<http://www.bipm.org/en/scientific/tai/>)

k	Institution
AMC	Alternate Master Clock station, Colorado Springs, CO, USA
AOS	Astrogeodynamical Observatory, Space Research Center P.A.S., Borowiec, Poland
APL	Applied Physics Laboratory, Laurel, MD, USA
AUS	Consortium of laboratories in Australia
BEV	Bundesamt für Eich- und Vermessungswesen, Vienna, Austria
BIM	Bulgarian Institute of Metrology, Sofiya, Bulgaria, formerly NMC
BIRM	Beijing Institute of Radio Metrology and Measurement, Beijing, P.R. China
BY	Belarussian State Institute of Metrology, Minsk, Belarus
CAO	Stazione Astronomica di Cagliari (Cagliari Astronomical Observatory), Cagliari, Italy
CH	Swiss Federal Office of Metrology, Switzerland (METAS)
CNM	Centro Nacional de Metrología, Querétaro, Mexico (CENAM)
CNMP	Centro Nacional de Metrología, de Panamá, Panamá
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center), Oberpfaffenhofen, Germany
DTAG	Deutsche Telekom AG, Frankfurt/Main, Germany
EIM	Hellenic Institute of Metrology, Thessaloniki, Greece
F	Commission Nationale de l'Heure, Paris, France
GUM	Główny Urząd Miar (Central Office of Measures), Warsaw, Poland
HKO	Hong Kong Observatory, Hong Kong, China
IFAG	Bundesamt für Kartographie und Geodäsie (Federal Agency for Cartography and Geodesy), Fundamental station, Wettzell, Kötzing, Germany
IGMA	Instituto Geográfico Militar, Buenos Aires, Argentina
INPL	National Physical Laboratory, Jerusalem, Israel
IPQ	Instituto Português da Qualidade, Monte de Caparica, Portugal

Table 13.1 Continued

k	Institution
IT	Istituto Nazionale di Ricerca Metrologica (INRIM), Italy
JATC	Joint Atomic Time Commission, Lintong, P.R. China
JV	Justervesenet, Norwegian Metrology and Accreditation Service, Kjeller, Norway
KIM	Research Center for Calibration, Instrumentation and Metrology The Indonesian Institute of Sciences, Serpong-Tangerang, Indonesia
KRIS	Korea Research Institute of Standards and Science, Daejeon, Rep. of Korea
MIKE	Center for Metrology and Accreditation, Finland
MKEH	Hungarian Trade Licensing Office, Hungary
LDS	University of Leeds, Leeds, United Kingdom
LT	Lithuanian National Metrology Institute, Vilnius, Lithuania
LV SA	Latvian National Metrology Center, Riga, Latvia
MSL	Measurement Standards Laboratory, Lower Hutt, New Zealand
NAO	National Astronomical Observatory, Misuzawa, Japan
NICT	National Institute of Information and Communications Technology, Tokyo, Japan
NIM	National Institute of Metrology, Beijing, P.R. China
NIMB	National Institute of Metrology, Bucharest, Romania
NIMT	National Institute of Metrology, Bangkok, Thailand
NIS	National Institute for Standards, Cairo, Egypt
NIST	National Institute of Standards and Technology, Boulder, CO, USA
NMIJ	National Metrology Institute of Japan, Tsukuba, Japan
NML	National Measurement Laboratory, Sydney, Australia
NMLS	National Metrology Laboratory of SIRIM Berhad, Shah Alam, Malaysia
NPL	National Physical Laboratory, Teddington, United Kingdom
NPLI	National Physical Laboratory, New Delhi, India
NRC	National Research Council of Canada, Ottawa, Canada
NTSC	National Time Service Center of China, Lintong, P.R. China

Table 13.1 *Continued*

k	Institution
ONBA	Observatorio Naval, Buenos Aires, Argentina
ONRJ	Observatório Nacional, Rio de Janeiro, Brazil
OP	Observatoire de Paris (Paris Observatory), Paris, France
ORB	Observatoire Royal de Belgique (Royal Observatory of Belgium), Brussels, Belgium
PL	Consortium of Laboratories in Poland
PTB	Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
ROA	Real Instituto y Observatorio de la Armada, San Fernando, Spain
SCL	Standards and Calibration Laboratory, Hong Kong
SG	Standards, Productivity and Innovation Board, Singapore (SPRING)
SIQ	Slovenian Institute of Quality and Metrology, Ljubljana, Slovenia
SMU	Slovenský metrology kústav (Slovak Institute of Metrology), Bratislava, Slovakia
SP	Sveriges Provnings-och Forskningsinstitut (Swedish National Testing and Research Institute), Borås, Sweden
SU	Institute of Metrology for Time and Space (IMVP), NPO 'VNIIFTRI' Mendeleevo, Moscow Region, Russia
SYRT	Laboratoire national de métrologie et d'essais – Système de Références Temps-Espace
TCC	TIGO Concepción Chile, Chile
TL	Telecommunication Laboratories, Chung-Li, Taiwan
TP	Institute of Photonics and Electronics, Czech Academy of Sciences, Praha, Czech Republic
UME	Ulusai Metroloji Enstitüsü, Marmara Research Center (National Metrology Institute), Gebze Kocaeli, Turkey
USNO	US Naval Observatory, Washington D.C., USA
VMI	Vietnam Metrology Institute, Ha Noi, Vietnam.
VSL	NMi Van Swinden Laboratorium, Delft, the Netherlands
ZA	Council for Scientific and Industrial Research, South Africa
ZMDM	Bureau of Measures and Precious Metals, Belgrade, Serbia

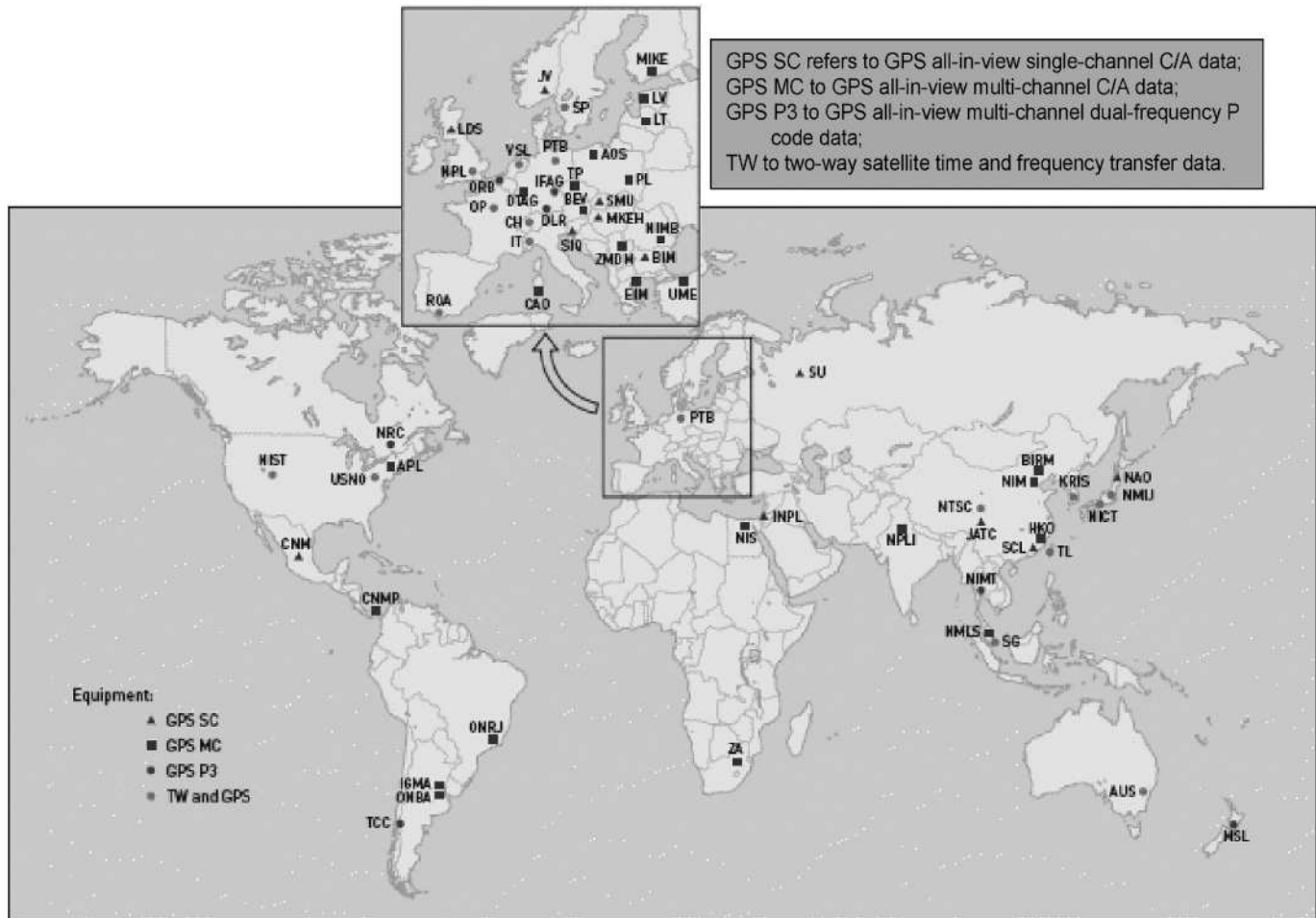


Figure 13.1 Distribution of institutions contributing to TAI (<http://www.bipm.org/en/scientific/tai/tai.html>).

parisons were to be done by portable clocks, the coordinate time accumulated during the transport of the portable clock from one point to another is given by:

$$\Delta t = \frac{2\omega}{c^2} A_E + \int_{\text{trajectory}} \left[1 - \frac{\Delta U(\mathbf{r})}{c^2} + \frac{\mathbf{v}}{2c^2} \right] ds, \quad (13.3)$$

where \mathbf{r} is a vector directed to the clock from the Earth's center; ds is the proper time element given by the clock; $\Delta U(\mathbf{r})$ is the difference in the gravity potential between the clock and the geoid (positive above the geoid); c is the speed of light; \mathbf{v} is the velocity of the clock relative to the Earth; ω is the angular speed of the Earth's rotation; and A_E is the equatorial projection of the area swept by \mathbf{r} in a coordinate system attached to the Earth.

$$A_E = \frac{1}{2} \int_{\text{trajectory}} \mathbf{r} \cdot \mathbf{v}_E \cos \phi \, ds, \quad (13.4)$$

\mathbf{v}_E being the eastward component of \mathbf{v} , and ϕ the geocentric latitude. The element of area is considered positive when the projection of \mathbf{r} rotates eastward. If the height of the clock above the surface of the Earth does not exceed 24 km, then it is possible to use $\Delta U(\mathbf{r}) = gh$, where g is the acceleration of gravity and h is the height of the clock above the geoid.

If the time comparison is done by means of an electromagnetic signal as opposed to a portable clock, the explanatory note called for the coordinate time elapsed between emission and reception of the signal to be given by

$$\Delta t = \frac{2\omega}{c^2} A_E + \frac{1}{c} \int_{\text{path}} d\sigma, \quad (13.5)$$

where $d\sigma$ is an increment of the proper length for the transmission.

The uncertainty of these comparisons is made up of two components. Type A errors describe the statistical uncertainty of frequency measures, that is, without any regard for the possible effect of calibration errors in the measurement. These errors currently range from ± 0.5 to ± 5.0 ns. The Type B error describes the systematic uncertainty of the calibration. Type B errors generally dominate, with estimates ranging from ± 1.0 to ± 20.0 ns. This situation could be improved with improved calibration techniques and more frequent calibration of the time transfer links.

Each clock H_i is assigned a statistical weight p_i based on its past stability. Its reading is also corrected by an expression:

$$h'_i(t) = a_i(t_0) + b_i(t)(t - t_0), \quad (13.6)$$

where the values of a_i and b_i are computed from previous analyses of the performance of the clock. EAL for any date, t , then is given by

$$\text{EAL}(t) = \frac{\sum_{i=1}^N p_i [h_i(t) + h'_i(t)]}{\sum_{i=1}^N p_i}, \quad (13.7)$$

where N is the number of clocks in the system.

These equations lead to the system (Tavella and Thomas, 1991)

$$\begin{aligned} x_{ij}(t) &= x_i(t) - x_j(t) \\ \sum_1^N p_i x_i(t) &= \sum p_i h'_i(t). \end{aligned} \quad (13.8)$$

The N clocks provide $N - 1$ difference equations, and the additional equation provided by the definition of EAL (13.7) allows the system to be solved exactly for each time t . The weights are chosen in an iterative procedure. In the first iteration the weights used are from the final iteration of the last published solution for EAL. In subsequent iterations the weights used are based on the statistics of the previous iteration. In each iteration, new values are calculated for the clock correction terms a_i and b_i in (13.6). The clock weights are then recalculated based on the reciprocal of the variance of the estimates of the clock's recent history of b_i corrections. To prevent any one clock, or any one set of clocks, from dominating the solution, a maximum weight limit is imposed. Also, in each iteration, every clock is checked for possible abnormal behavior, and, if a clock fails a statistical test for normal behavior, its weight is set to zero. Each iteration provides values of x_i , b_i , and p_i for every clock in the solution (Thomas and Azoubib, 1996).

Figure 13.2 shows the variation in the number of clocks contributing to the formation of EAL from 1999 to 2007, and Figure 13.3 shows the stability of EAL from 1999 to 2006.

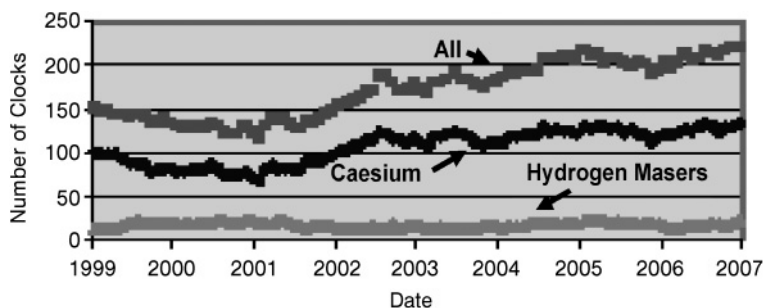


Figure 13.2 Number of clocks contributing to EAL from Petit (2007).

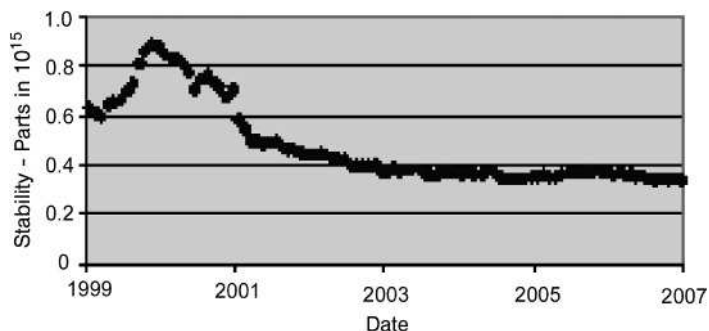


Figure 13.3 Stability of EAL from Petit (2007).

Since its inception, the stability of EAL has improved, mainly because of improvements in the contributing clocks and changes in the weighting schemes that were adjusted to take advantage of the improvement in the standards. After January 1998 the maximum weight was set to a fixed value; after January, 2001, that maximum was set to $2/N$, where N is the number of clocks used; after July, 2002, that maximum weight was adjusted to $2.5/N$ (Petit, 2003).

13.3.2

Steering EAL with Primary Frequency Standards

The requirement for accuracy (as opposed to stability) in a time scale is met by making use of data from specific primary laboratory standards to provide systematic corrections to the scale interval of EAL. The primary standards are frequency standards maintained at a few laboratories that realize the second with the highest accuracy. These institutions make corrections for various systematic effects and are consequently able to provide an accurate realization of the second with uncertainties on the order of 1×10^{-16} . Laboratories currently providing primary frequency standard data include IT, NIST, NMIJ, NPL, PTB, and SYRTE. The data supplied to the BIPM for each frequency comparison include the interval over which the comparison was done, the uncertainty of the measurement due to systematic effects in the primary frequency standard, the uncertainty due to the instability of the primary frequency standard, the uncertainty due to the link between the primary frequency standard and the clock whose data were used in the EAL calculation, the uncertainty in the link between that clock and the EAL reference, and a reference publication dealing with the evaluation of the systematic error (Petit, 2000). Corrections are made to account for the relativistic frequency shift caused by the distance of the standards above or below a conventional surface of equal gravity potential, very close to the rotating geoid. This correction is about 1×10^{-16} per meter. The total fractional frequency uncertainties of these data, including the errors due to the frequency transfers, range from ± 0.5 to $\pm 12 \times 10^{-15}$.

The BIPM uses these data in an algorithm created originally at the BIH to produce the steering corrections to the EAL (Azoubib, Granveaud and Guinot, 1977). The combination of EAL with these systematic corrections provides the final product, TAI. Before 1 January 1977 TAI was equal to EAL. After that date a systematic steering correction of 1×10^{-12} was applied to ensure agreement between the TAI scale unit and the SI second at sea level. After that, EAL was modified using steps of 2×10^{-14} as needed to bring the scale unit into agreement with the SI second as defined by the primary frequency standards (Guinot, 1988). Following the CCDS recommendation S2 of 1996 (Quinn, 1997), the frequencies of the primary frequency standards were corrected systematically for black body radiation. From 1998 to 2004, steering corrections of $\pm 1 \times 10^{-15}$ were applied as needed, for intervals of 2 months at least. After July, 2004, frequency corrections, up to 0.7×10^{-15} , are applied for intervals of 1 month, if the value of the difference between the TAI scale unit and the SI second provided by the primary frequency

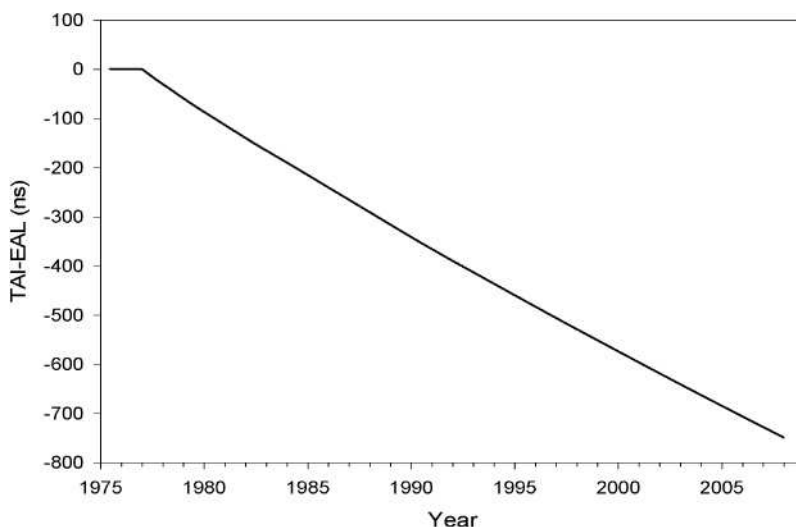


Figure 13.4 TAI-EAL.

standards exceeds 2.5 times its uncertainty (Guinot and Arias, 2005). The SI Brochure Appendix 2 – Practical realization of the definition of the unit of time (BIPM, 2006) reports that the difference between the TAI scale unit and the SI second on the rotating geoid was a few parts in 10^{15} in 2006 with an uncertainty of 10^{-15} . It also reports typical steering corrections to TAI of a few parts in 10^{16} per month, which are reported in the monthly BIPM Circular T available at <http://www.bipm.org/en/scientific/tai/>. Figure 13.4 shows the difference between the EAL and TAI time scales since 1975.

13.4

Stability of TAI

The current estimate of the stability of TAI created using the process outlined above is 0.5×10^{-15} for averaging times of 20–40 days (Guinot and Arias, 2005). The long-term stability is limited by the accuracy of the measurements of the primary frequency standards, which is currently at the level of 10^{-15} . Figure 13.5 shows the stability of EAL from 2000 through 2006 and TAI for two different periods: 1993–1999 and 2000–2006 as a comparison to Terrestrial Time (TT).

13.5

Distribution of TAI

Figure 13.6 displays the flow of data connected to the formation of TAI.

TAI is available to users with a typical delay of a few weeks. However, it is not distributed directly, but is made available through the use of Coordinated Universal

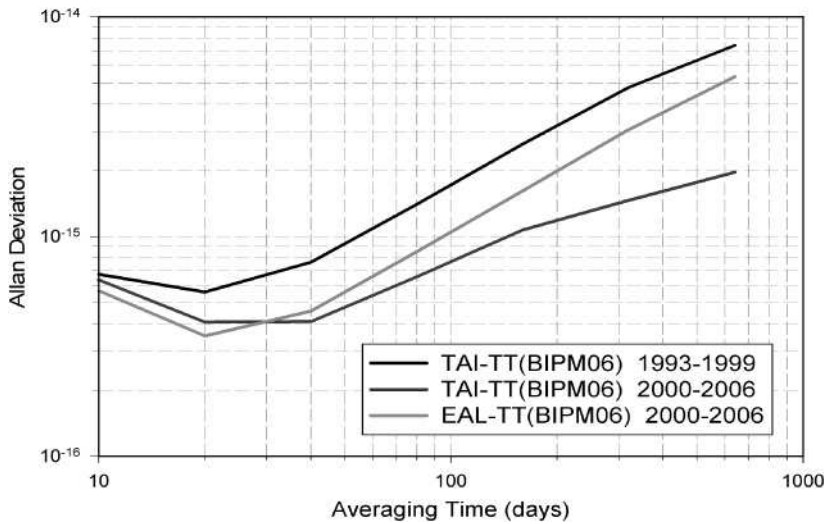


Figure 13.5 Stability of TAI and EAL.

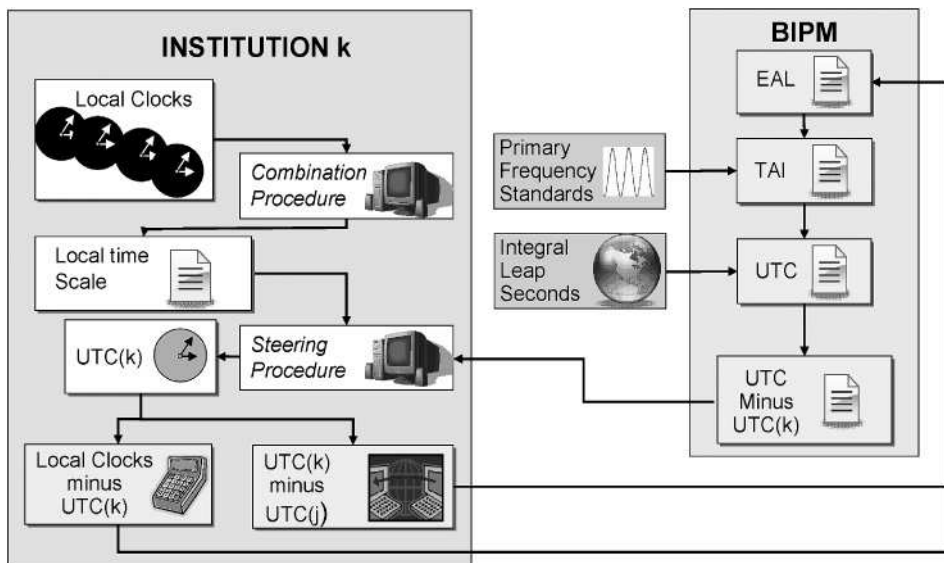


Figure 13.6 Data flow in the formation of TAI.

Time (UTC) (see Chapter 14). UTC differs from TAI only by an integral number of seconds, so it is clear that $\text{UTC} - h_i(t)$ is equivalent to $\text{TAI} - h_i(t)$ modulo integral seconds, $h_i(t)$ being the contributed clock readings (see Section 13.1). UTC is used formally to provide this accessibility as it is the conventionally accepted international standard and the time scale in common use. Institutions that contribute to the formation of TAI by the BIPM maintain their local approximation, designated

UTC(k), of the international standard. Here the letter k designates an institutional abbreviation (see Table 13.1). The conventional nomenclature stipulates that UTC and TAI, appearing without a following set of parentheses, indicates the scale that is provided by the BIPM. Monthly the BIPM disseminates the time series UTC–UTC(k) at five-day intervals, and since 2005 they also provide the uncertainties of these offsets. From this information the contributing laboratories can then access TAI, since TAI and UTC differ only by integral seconds (BIPM, 2006).

This information appears in the monthly publication *Circular T*. This periodical bulletin also contains the data by which the contributing institutions can steer their respective clocks to match the frequency of TAI to meet their respective requirements. It provides access to the realization of the SI second by listing the deviation of the TAI scale interval with respect to the SI second, and the individual observations of each of the contributing primary frequency standards. The time links used in the calculation of TAI and their uncertainties are also shown. All data used to calculate TAI and *Circular T*, as well as related information, can be found on the ftp server of the BIPM time section (www.bipm.org).

13.6

Relationship of TAI to Terrestrial Time

Since its official beginning in 1972 with 56 clocks and 25 laboratories, TAI has been maintained as a continuous time scale that can be considered as a realization of Terrestrial Time (TT). TT was defined by the IAU in a 1991 resolution:

The XXth General Assembly of the International Astronomical Union,

Considering

1. that the time scales used for dating events observed from the surface of the Earth and for terrestrial metrology should have as the unit of measurement the SI second, as realized by terrestrial time standards,
2. the definition of the International Atomic Time, TAI, approved by the 14th Conférence Générale des Poids et Mesures (1971) and completed by a declaration of the 9th session of the Comité Consultatif pour la Définition de la Seconde (1980).

Recommends that,

1. the time reference for apparent geocentric ephemerides be Terrestrial Time, TT,
2. TT be a time scale differing from TCG [Geocentric Coordinate Time] of Recommendation III by a constant rate, the unit of measurement of TT being chosen so that it agrees with the SI second on the geoid,

3. at instant 1977 January 1, 0h 0m 0s TAI exactly, TT have the reading 1977 January 1, 0h 0m 32.184s exactly.

Notes for Recommendation IV

1. The basis of the measurement of time on the Earth is International Atomic Time (TAI), which is made available by the dissemination of corrections to be added to the readings of national time scales and clocks. The time scale TAI was defined by the 59th session of the Comité International des Poids et Mesures (1970) and approved by the 14th Conférence Générale des Poids et Mesures (1971) as a realized time scale. As the errors in the realization of TAI are not always negligible, it has been found necessary to define an ideal form of TAI, apart from the 32.184s offset, now designated Terrestrial Time, TT.
2. The time scale TAI is established and disseminated according to the principle of coordinate synchronization, in the geocentric coordinate system, as explained in CCDS, 9th Session (1980) and in Reports of the CCIR, 1990, annex to Volume VII (1990).
3. In order to define TT it is necessary to define the coordinate system precisely, by the metric form to which it belongs. To be consistent with the uncertainties of the frequency of the best standards, it is at present (1991) sufficient to use the relativistic metric given in Recommendation I.
4. For ensuring an approximate continuity with the previous time arguments of ephemerides (Ephemeris Time, ET), a time offset is introduced so that $TT - TAI = 32.184s$ exactly at 1977 January 1, 0h TAI. This date corresponds to the implementation of a steering process of the TAI frequency, introduced so that the TAI unit of measurement remains in close agreement with the best realizations of the SI second on the geoid. TT can be considered as equivalent to TDT as defined by International Astronomical Union (IAU) Recommendation 5 (1976) of Commissions 4, 8 and 31, and Recommendation 5 (1979) of Commissions 4, 19 and 31.
5. The divergence between TAI and TT is a consequence of the physical defects of atomic time standards. In the interval 1977–1990, in addition to the constant offset of 32.184s, the deviation probably remained within the approximate limits of ± 10 microseconds. It is expected to increase more slowly in the future as a consequence of improvements in time standards. In many cases, especially for the publication of ephemerides, this deviation is negligible. In such cases, it can be stated that the argument of the ephemerides is $TAI + 32.184s$.
6. Terrestrial Time differs from TCG of Recommendation III by a scaling factor, in seconds:

$$TCG - TT = L_G \times (JD - 2\,443\,144.5) \times 86\,400.$$

The present estimate of the value of L_G is $6.969\,291 \times 10^{-10}$ ($\pm 3 \times 10^{-16}$). The numerical value is derived from the latest estimate of gravitational potential on the

geoid, $W = 62\,636\,860 (+/-30) \text{ m}^2 \text{ s}^{-2}$ (Chovitz, *Bulletin Géodésique*, **62**, 359, 1988). The two time scales are distinguished by different names to avoid scaling errors. The relationship between L_B and L_C of Recommendation III, notes 1 and 2, and L_G is, $L_B = L_C + L_G$.

7. The unit of measurement of TT is the SI second on the geoid. The usual multiples, such as the TT day of 86 400 SI seconds on the geoid and the TT Julian century of 36 525 TT days, can be used, provided that the reference to TT be clearly indicated whenever ambiguity may arise. Corresponding time intervals of TAI are in agreement with the TT intervals within the uncertainties of the primary atomic standards (e.g., within $\pm 2 \times 10^{-14}$ in relative value during 1990).
8. Markers of the TT scale can follow any date system based upon the second, for example, the usual calendar date or the Julian Date, provided that the reference to TT be clearly indicated whenever ambiguity may arise.
9. It is suggested that realizations of TT be designated by TT(yyy) where yyy is an identifier. In most cases a convenient approximation is:

$$TT(TAI) = TAI + 32.184 \text{ s.}$$

However, in some applications it may be advantageous to use other realizations. The BIPM, for example, has issued time scales such as TT(BIPM90).

In 2000 the IAU re-defined TT with the following resolution:

Resolution B1.9 Re-definition of Terrestrial Time TT

The XXIVth International Astronomical Union General Assembly,

Considering

1. that IAU Resolution A4 (1991) has defined Terrestrial Time (TT) in its Recommendation 4,
2. that the intricacy and temporal changes inherent to the definition and realization of the geoid are a source of uncertainty in the definition and realization of TT, which may become, in the near future, the dominant source of uncertainty in realizing TT from atomic clocks,

Recommends

that TT be a time scale differing from TCG by a constant rate: $dTT/dTCG = 1 - L_G$, where $L_G = 6.969\,290\,134 \times 10^{-10}$ is a defining constant,

Note

L_G was defined by the IAU Resolution A4 (1991) in its Recommendation 4 as equal to U_G/c^2 where U_G is the geopotential at the geoid. L_G is now used as a defining constant.

These resolutions make it clear that, for practical applications, TAI can be considered a near real-time realization of TT and that effectively $TT = TAI + 32.184\text{ s}$. In practice TT is made available annually after a re-analysis of the data used to produce TAI. These versions of TT are designated TT(BIPMyy) where the yy stands for the last two digits of the last year of data used in the reanalysis. Thus TT(BIPM06), for example, was computed in January, 2007, using all of the data contributed through 2006. Figure 13.7 shows the estimated accuracy of TT(BIPM06).

TT(BIPM07) is the most recent version available from the BIPM at <ftp://ftp2.bipm.org/pub/tai/scale/TTBIPM/ttbipm.07>. It is identical to all past realizations since TT(BIPM99) for all dates before 2 January 1993 and is within 1 ns of the previous realization TT(BIPM06) for dates before September 2006. The difference between these two versions of TT reaches 2 ns in December 2006. Figure 13.8 shows a plot of the difference between TT(BIPM07) and TAI.

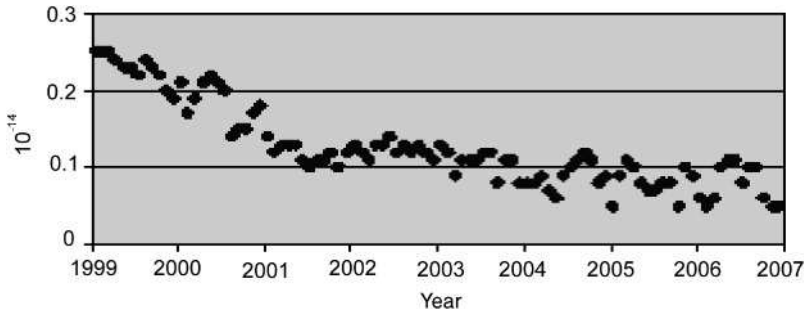


Figure 13.7 Estimated accuracy of TT(BIPM06) from Petit (2007).

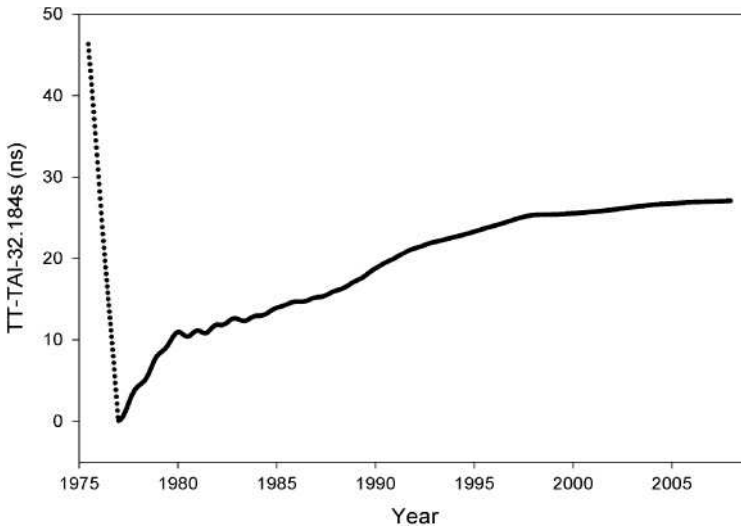


Figure 13.8 Difference between TT(BIPM07) and TAI.

References

- Audoin, C. and Guinot, B. (2001) *The Measurement of Time*, Cambridge University Press, Cambridge, p. 335.
- Azoubib, J., Granveaud, M. and Guinot, B. (1977) Estimation of the scale unit duration of time scales. *Metrologia*, **13**, 87–93.
- Barnes, J.A., Andrews, D.H. and Allan, D.W. (1965) The NBS-A time scale—its generation and dissemination. *IEEE Trans. Instrum. Meas.*, **IM-14**, 228–32.
- Bureau International de l'Heure (1974) *BIH Annual Report for 1973*, Observatoire de Paris (in French; English version appears in *BIH Annual Report for 1974*, published in 1975).
- Bureau International des Poids et Mesures (2006) *The International System of Units (SI)*, Eighth Edition, Bureau International des Poids et Mesures.
- Comptes Rendus de la 14e CGPM (1971), 1972, 78. Available at <http://www1.bipm.org/en/convention/cgpm/resolutions.html>.
- Comptes Rendus de la 15e CGPM (1975), 1976, 104. Available at <http://www1.bipm.org/en/convention/cgpm/resolutions.html>.
- Comptes Rendus de la 18e CGPM (1987), 1988, 98. Available at <http://www1.bipm.org/en/convention/cgpm/resolutions.html>.
- Essen, L. *Time for Reflection*, published privately and available at <http://www.btinternet.com/~time.lord/> also available in Henderson, D. (2005) Essen and the National Physical Laboratory's atomic clock. *Metrologia*, **42**, S4–9.
- Essen, L. and Parry, J.V.L. (1957) The caesium resonator as a standard of frequency and time. *Philos. Trans. Roy. Soc. London. Ser. A, Math. Phys. Sci.*, **250**, 45–69.
- Giacomo, P. (1981) News from the BIPM. *Metrologia*, **17**, 69–74.
- Giacomo, P. (1988) News from the BIPM. *Metrologia*, **25**, 113–19.
- Guinot, B. (1988) Atomic time scales for pulsar studies and other demanding applications. *Astron. Astrophys*, **192**, 370–3.
- Guinot, B. and Arias, E.F. (2005) Atomic time-keeping from 1955 to the present. *Metrologia*, **42**, S20–30.
- Markowitz, W. (1962) *IRE Trans. Instrum.*, **I-11**, 239–42.
- Nelson, R.A., McCarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L. and Bartholomew, T.R. (2001) The leap second: its history and possible future. *Metrologia*, **38**, 509–29.
- Petit, G. (2000) Use of primary frequency standards for estimating the duration of the scale unit of TAI. *31st Annual Precise Time and Time Interval (PTTI) Meeting*, pp. 297–304.
- Petit, G. (2003) A new realization of terrestrial time. *35th Annual Precise Time and Time Interval (PTTI) Meeting*, pp. 307–18.
- Petit, G. (2007) The long term stability of EAL and TAI (revisited). *Frequency Control Symp. 2007 Joint with the 21st Euro Frequency & Time Forum*, pp. 391–4.
- Quinn, T.J. (1997) News from the BIPM. *Metrologia*, **34**, 187–94.
- Tavella, P. and Thomas, C. (1991) Comparative study of time scale algorithms. *Metrologia*, **28**, 57–63.
- Terrien, J. (1970) News from the Bureau International des Poids et Mesures. *Metrologia*, **7**, 43–4.
- Terrien, J. (1971) News from the Bureau International des Poids et Mesures. *Metrologia*, **8**, 32–6.
- Terrien, J. (1975) News from the Bureau International des Poids et Mesures. *Metrologia*, **11**, 1789–183.
- Thomas, C. and Azoubib, J. (1996) TAI computation: study of an alternative choice for implementing an upper limit of clock weights. *Metrologia*, **33**, 227–40.
- Time Service Notice No. 6 (1 January 1959) U.S. Naval Observatory, Washington, D.C.
- Trans. Int. Astron. Union (1962) *Reports on Astronomy*, Vol. XI A (ed D.H. Sadler), Academic Press, New York, pp. 362–3.

14

Coordinated Universal Time

14.1

Universal Time Before 1972

In the nineteenth century the words ‘universal time’ were typically used to refer to the concept of time that would read the same everywhere in the world and be used as a conventional, or ‘universal,’ time standard. This is in contrast with the common practice involving many ‘local’ times referred to local meridians. The phrase did not necessarily refer to a particular time scale such as Greenwich Mean Time (GMT), the mean solar time of the Greenwich meridian in England (see Chapter 2 for the definitions and history of GMT). The most precise determinations of time in the nineteenth century were accomplished using astronomical observations of star transits of the observers’ meridians. Consequently, the development of the concept of universal time is related directly to the acceptance of the idea of a standard meridian to which those astronomical observations could be referred.

The commencement of the British *Nautical Almanac* in 1767 had predisposed users of maritime charts to the Greenwich meridian. However, charts based on many other ‘standard’ meridians were available including Christiania (Oslo), Copenhagen, Naples, Paris, and Stockholm. In August, 1871, the first International Geographical Congress met in Antwerp and passed a resolution expressing the participants’ opinion that the Greenwich meridian should be used as the zero of longitude for all passage charts and that this should be obligatory within 15 years (Howse, 1997). The second International Geographical Congress met in Rome in 1875, producing further discussion without definite results. At this meeting, however, the proposition was first suggested that France might consider adopting the Greenwich meridian if Great Britain were to adopt the metric system.

In 1878 Sandford Fleming, the engineer-in-chief of the Canadian Pacific Railway published an article that promoted the concept of a universal time (Fleming, 1878). This was followed in 1879 by two papers outlining his ideas regarding time (Fleming, 1879a, 1879b). In these works he proposed what he first called ‘cosmo-



Figure 14.1 Cosmopolitan time clock face (Fleming, 1879a).

politan' time. However, he states that 'For this purpose either of the designations, "common," "universal," "nonlocal," "uniform," "absolute," "all world," "terrestrial," or "cosmopolitan," might be employed.' (Fleming, 1879a). The words 'cosmic time' were also used. The globe would be divided into 24 separate lunes, each corresponding to a chosen meridian of longitude designated by a letter of the alphabet, and cosmopolitan time would correspond to the time of the initial or prime meridian. To distinguish cosmopolitan time from a locally realized time he suggested the use of a 24-hour system where the hours would be distinguished by letters (Figure 14.1). Although Fleming suggested the use of a prime meridian, he did not propose the use of Greenwich, because he apparently felt that this would be too politically sensitive (Blaise, 2002). He eventually favored the adoption of a meridian situated 180 degrees from that of Greenwich corresponding loosely to the current 'date line.'

In 1880 GMT did become the legal time in Great Britain, and in 1883 the U.S. and Canadian railways adopted a system of time zones based on the Greenwich meridian to facilitate scheduling. The U.S. government did not implement a time zone system officially until 1918 (Bartky, 2000). Meanwhile, the Third International Geographical Congress met in 1881 in Venice to discuss the zero meridian and a standard time, among other issues. The participants voted to appoint an international commission to consider the problem, but no action was taken (Wheeler, 1885). In 1883 the issues were taken up at the Seventh General Conference of the International Association of Geodesy held in Rome. There, the delegates adopted resolutions that, among other things, (i) suggested Greenwich as the initial meridian, (ii) recommended that longitude be measured from west to east, (iii) recognized the usefulness of adopting 'une heure universelle' in addition to 'heures locales,' (iv) recommended that Greenwich noon, which corresponds with midnight on the meridian situated 12 hours from Greenwich in longitude, be the beginning of the cosmopolitan date, and (v) noted the convenience of measuring time from 0h to 24h. They also noted the special conference that had been proposed by the U.S. government regarding the standardization of longitude and time (Hirsch and Von Oppolzer, 1883).

The International Meridian Conference held in Washington in October 1884 settled the matter by proposing 'the meridian passing through the center of the transit instrument at the Observatory of Greenwich as the initial meridian for



Figure 14.2 Greenwich prime meridian.

longitude' (Figure 14.2). Participants in that conference also took on the issue of an international convention for time by proposing '... the adoption of a universal day for all purposes for which it may be found convenient, and which shall not interfere with the use of local or other standard time where desirable.' They further proposed that '... this universal day is to be a mean solar day; is to begin at the moment of mean midnight of the international meridian, coinciding with the beginning of the civil day and date of that meridian; and is to be counted from zero up to twenty-four hours.' (*International Conference held at Washington for the purpose of fixing a Prime Meridian and a Universal Day, October 1884 – Protocols of the Proceedings*).

Despite the recommendations of the 1884 International Meridian Conference, astronomers continued to measure days from noon to noon. Following that tradition, the mean solar time measured from mean noon at Greenwich was designated as Greenwich Mean Time (GMT). In 1919 the Bureau International de l'Heure (BIH), the international service for time, began to coordinate the emission of time signals by radio stations based on Greenwich Civil Time (GCT), which is GMT plus 12h, following the recommendation of the International Meridian Conference. In 1925, however, the situation was changed in the astronomical almanacs by introducing a 12-hour discontinuity whereby the date previously referred to as 31.5 December 1924 was now to be known as 1.0 January 1925. The *British Nautical Almanac* continued to call this time Greenwich Mean Time, but *The American Ephemeris* referred to this new time scale, measured from midnight to midnight, as Greenwich Civil Time. To avoid confusion, the name 'Greenwich Mean Astronomical Time' (GMAT) was used to designate the time measured from noon to noon.

In 1928, The IAU recommended using the name 'Universal Time' to replace GMT or GCT in astronomical almanacs. This was the first 'official' designation of Universal Time. The International Research Council had established the Bureau International de l'Heure (BIH) at the Paris Observatory in 1919 to coordinate the transmission of radio time signals. It published routinely the difference between the broadcast radio signal and the astronomically determined time.

The actual determination of this time continued to rely on astronomical observations of star transits that were used to set mechanical, and later, electronic clocks. Beginning in 1956 the IAU recognized three versions of Universal Time. The Greenwich Mean Solar Time as observed at any location on the Earth, without regard for the location of the Earth's rotation axis with respect to the observing site, was designated 'UT0.' If we also know the position of the pole with respect to the observing location we can apply small corrections (on the order of tens of milliseconds) to produce a time scale, 'UT1,' that is free of the local effects of the station's geography. Finally, a third version was designated 'UT2,' that was obtained by applying a conventionally adopted seasonal variation to UT1 to account for the observed seasonal variation in the Earth's rotational speed. This time was generally regarded in the early 1950s as being the best representation of a uniform time scale, and radio time signals of that time were based on UT2.

In 1944 quartz crystal clocks began to be used to broadcast time signals. These devices kept time with a uniform rate and were adjusted as needed to keep pace with time determined astronomically. Atomic clocks based on the frequency of an atomic transition in the cesium atom became available in 1955. A radio station broadcasting the national time standard for the UK began sending time signals, determined using an atomic clock based on a provisional calibration of the frequency of the atomic transition. In the United States, the U.S. Naval Observatory and the National Bureau of Standards (NBS) (now the National Institute of Standards and Technology) also developed time scales based on cesium atomic clocks. This work was done from 1956 to 1957 and, as a result, the NBS radio station WWV began broadcasting time signals based on atomic clocks that were adjusted in rate and offsets to match the UT2 that was determined from star transits.

The World Administrative Radio Congress of 1959 recognized that different countries were sending inconsistent time signals, and they asked the International Radio Consultative Committee, abbreviated 'CCIR' to study the problem. The United Kingdom and the United States had decided in 1957 to combine Nautical Almanacs beginning with the 1960 edition, and in 1959 they also agreed to coordinate their time and frequency transmissions by making the same adjustments, at the same time, to their cesium-based time scales to stay close to UT2. In 1959 the Royal Greenwich Observatory, the National Physical Laboratory in England, and the U.S. Naval Observatory agreed to coordinate their time and frequency transmissions, which were based on UT2 and the atomic frequency. Based on a comparison of UT2 and the rotation rate of the Earth during the previous year, a factor S was determined, and the actual frequency of transmission would be F_0

$(1 + S)$, where F_0 is the nominal atomic frequency. The time between pulses was $9\,192\,631\,770 (1 - S)$ cycles of the cesium resonance. When the rotation of the Earth departed unpredictably from this offset atomic scale, step adjustments were introduced in the time scale in multiples of 50 milliseconds. The purpose of this cooperation was to avoid diverse time scales and to provide the same time and frequency from multiple sources. This coordination began on January 1, 1960, and the resulting time scale began to be called informally ‘Coordinated Universal Time.’ Timing laboratories from other countries also began to participate over time, and in 1961 the Bureau International de l’Heure at Paris Observatory began to coordinate the process internationally.

The original form of UTC was formalized in CCIR Recommendation 374 in 1963. However, the National Bureau of Standards (NBS) continued to refer to their time signals as GMT. In 1965 the BIH started calculating UTC based on the atomic time scale A3 that would eventually evolve into TAI. Each year, the BIH would, after consulting other observatories, announce an offset in the atomic frequency in order to match UT2 as closely as possible. They would also announce 100 ms adjustments in UTC as required in order to maintain UTC with 0.1 s of UT2. In 1967 the CCIR adopted the names *Coordinated Universal Time* and *Temps Universel Coordonné* for the English and French names with the acronym UTC to be used in both languages. The name ‘Coordinated Universal Time (UTC)’ was approved by a resolution of IAU Commissions 4 and 31 at the 13th General Assembly in 1967 (Trans. Int. Astron. Union, 1968). For provisional limited use, the CCIR in 1966 approved ‘Stepped Atomic Time,’ which used the atomic second with frequent 200-ms adjustments made in order to be within 0.1 s of UT2.

The resulting UTC time scale broadcast worldwide, with its seconds of variable length and potential ‘jumps’ in time, began to cause concerns among users that needed stable time scales. There was an increasing need for precise frequencies for both military and civilian applications. Radio and television stations needed precise frequency standards to calibrate their transmitters so they could stay within their assigned places in the overcrowded frequency spectrum. Precise calibration of oscillators was also required for navigation systems such as Loran, Loran C, and Omega. Thus, the changing offset frequency was becoming a nuisance, and an attempt was made to maintain the same frequency for several years at a time. The proposed introduction of an air collision avoidance system in the early 1970s, based on precise frequency, made the use of frequency offsets intolerable. These concerns drove the acceptance of a new UTC, adopted in 1970 and implemented in 1972.

14.2

Coordinated Universal Time After 1972

The current system of Coordinated Universal Time can be traced back to a meeting of the International Union of Radio Science (URSI) in 1966 when participants noted the need for a uniform atomic frequency. At the 1967 meeting of URSI,

participants agreed that all adjustments to atomic time should be eliminated, and that UT2 information could be distributed in tables or in radio transmissions. In May, 1968, the idea of the current practice of introducing leap second adjustments in the UTC time scale was introduced independently by Louis Essen and Gernot Winkler at a meeting of a commission organized by the International Committee for Weights and Measures (CIPM) to discuss the issue. In that same year, CCIR Study Group 7, meeting in Boulder, Colorado, discussed possible changes in the definition of UTC Nelson *et al.*, 2001).

They formed an 'Interim Working Party' to provide proposals for a possible new definition of UTC. The options considered were (i) steps in UTC of 0.1 or 0.2 seconds to keep UTC close to UT2, (ii) replacing UTC with a time scale with no adjustments, and (iii) one-second adjustments.

Study Group 7 then formulated specific proposals that were approved in January 1970 at the CCIR XIIth Plenary Assembly in New Delhi. The recommendation adopted there provides the current definition of the world's civil time. It specified that (i) radio carrier frequencies and time intervals should correspond to the atomic second based on the cesium atom; (ii) step adjustments should be exactly one second to maintain approximate agreement with UT; and (iii) standard time signals should contain information on the difference between UTC and UT. The new system was to begin on 1 January 1972. In February 1971 Study Group 7 specified more details regarding the implementation of the 1970 recommendation 460. The predicted difference $DUT1 = UT1 - UTC$ was to be coded into the broadcast time signals, and $DUT1$ was not to exceed 0.7 s. A special offset of -0.1077580 second was given to UTC at the end of 1971, so that $TAI - UTC$ was exactly 10 seconds. Since then the UTC scale has been based on TAI with leap seconds added to keep UTC within less than a second of UT1.

The CCIR failed to send an official letter concerning the change to the IAU in time for its 1970 General Assembly. Hence, the IAU could not respond until the 1973 General Assembly, which was after the introduction of the change. In 1973 the IAU recognized that UTC provided mean solar time, recommended it for civil time, and suggested modifications to the leap second rules. In 1974 the CCIR revised recommendation 460-1 based on the input from the IAU, and raised the maximum difference between UTC and UT1 to 0.9 second. In 1975 the CGPM stated that UTC provided both atomic frequency and UT, and endorsed it for civil time. On 1979 January 1 the rate of TAI was reduced by one part in 10^{12} to better approximate the SI second. Thus, the UTC rate was also changed. The CCTF determined that the TAI second was longer than the SI second, because the time standards were not being corrected for the effects of black body radiation. So from 1996 to 1998 the TAI was steered to reduce the length of the second by two parts in 10^{14} . This also had a corresponding effect on UTC.

In 1988 the responsibility for TAI was transferred to the BIPM from the BIH, and the responsibility for determining the rotation of the Earth and UT1 was

transferred to the IERS. Thus, both the BIH and the International Latitude Service (ILS) ceased to exist at that time. In a 1992 reorganization, the International Telecommunications Union-Radiocommunications Sector (ITU-R) replaced the CCIR, and the UTC recommendation became ITU-R TF 460.

The current UTC system is defined by ITU-R (formerly CCIR) Recommendation ITU-R TF.460-5:

UTC is the time scale maintained by the BIPM, with assistance from the IERS, which forms the basis of a coordinated dissemination of standard frequencies and time signals. 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.

14.3

Leap Seconds

To maintain the tolerance of 0.9 s between UTC and UT1, positive leap seconds have been introduced as needed since 1972. A positive, or negative, leap second could be the last second of any UTC month, but first preference should be given to the end of December and June, and second preference to the end of March and September. A positive leap second begins at 23 h 59 m 60 s and ends at 0 h 0 m 0 s of the first day of the following month. In that case the progression of seconds would read:

23 h 59 m 58 s
 23 h 59 m 59 s
 23 h 59 m 60 s
 24 h 00 m 00 s

In the case of a negative leap second, 23 h 59 m 58 s will be followed one second later by 0 h 0 m 0 s of the first day of the following month. The IERS should decide upon and announce the introduction of a leap second at least eight weeks in advance. The first leap second was introduced on 30 June 1972. On average leap seconds would be expected to occur approximately every 18 months at the present time, because of the secular deceleration of the Earth. However, because of other changes in the Earth's rotation, these events cannot be predicted far in advance. Astronomical observations of the Earth's rotation angle are required to determine when leap seconds should be inserted. Table 14.1 provides the mathematical relationship between UTC and TAI since 1961.

The IERS maintains updates to this table at www.iers.org. Figure 14.3 plots the difference between TAI and UT1 along with the difference between TAI and UTC to show the relationships among these scales since 1961.

Table 14.1 TAI-UTC.

From			To			TAI-UTC			
1961	Jan.	1	1961	Aug.	1	1.422 818 s	+	(MJD-37300)	× 0.001 296 s
	Aug.	1	1962	Jan.	1	1.372 818 s	+	(MJD-37300)	× 0.001 296 s
1962	Jan.	1	1963	Nov.	1	1.845 858 s	+	(MJD-37665)	× 0.001 123 2 s
1963	Nov.	1	1964	Jan.	1	1.945 858 s	+	(MJD-37665)	× 0.001 123 2 s
1964	Jan.	1		April	1	3.240 130 s	+	(MJD-38761)	× 0.001 296 s
	April	1		Sept.	1	3.340 130 s	+	(MJD-38761)	× 0.001 296 s
	Sept.	1	1965	Jan.	1	3.440 130 s	+	(MJD-38761)	× 0.001 296 s
1965	Jan.	1		Mar.	1	3.540 130 s	+	(MJD-38761)	× 0.001 296 s
	Mar.	1		Jul.	1	3.640 130 s	+	(MJD-38761)	× 0.001 296 s
	Jul.	1		Sept.	1	3.740 130 s	+	(MJD-38761)	× 0.001 296 s
	Sept.	1	1966	Jan.	1	3.840 130 s	+	(MJD-38761)	× 0.001 296 s
1966	Jan.	1	1968	Feb.	1	4.313 170 s	+	(MJD-39126)	× 0.002 592 s
1968	Feb.	1	1972	Jan.	1	4.213 170 s	+	(MJD-39126)	× 0.002 592 s
1972	Jan.	1		Jul.	1	10 s			
	Jul.	1	1973	Jan.	1	11 s			
1973	Jan.	1	1974	Jan.	1	12 s			
1974	Jan.	1	1975	Jan.	1	13 s			
1975	Jan.	1	1976	Jan.	1	14 s			
1976	Jan.	1	1977	Jan.	1	15 s			
1977	Jan.	1	1978	Jan.	1	16 s			
1978	Jan.	1	1979	Jan.	1	17 s			
1979	Jan.	1	1980	Jan.	1	18 s			
1980	Jan.	1	1981	Jul.	1	19 s			
1981	Jul.	1	1982	Jul.	1	20 s			
1982	Jul.	1	1983	Jul.	1	21 s			
1983	Jul.	1	1985	Jul.	1	22 s			
1985	Jul.	1	1988	Jan.	1	23 s			
1988	Jan.	1	1990	Jan.	1	24 s			
1990	Jan.	1	1991	Jan.	1	25 s			
1991	Jan.	1	1992	Jul.	1	26 s			
1992	Jul.	1	1993	Jul.	1	27 s			
1993	Jul.	1	1994	Jul.	1	28 s			
1994	Jul.	1	1996	Jan.	1	29 s			
1996	Jan.	1	1997	Jul.	1	30 s			
1997	Jul.	1	1999	Jan.	1	31 s			
1999	Jan.	1	2006	Jan.	1	32 s			
2006	Jan.	1	2009	Jan.	1	33 s			
2009	Jan.	1				34 s			

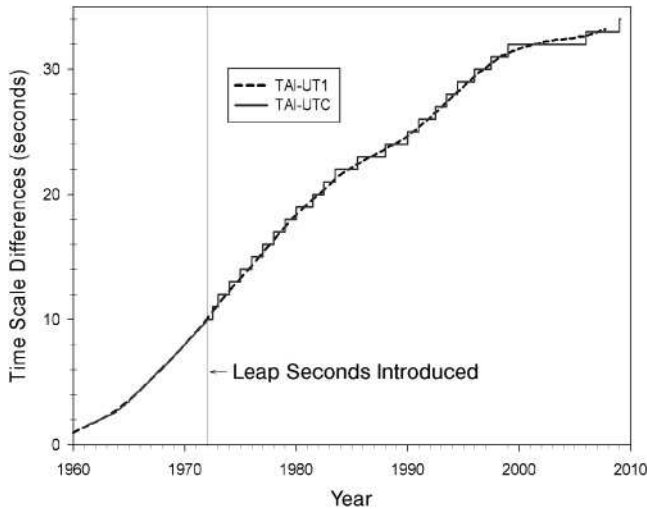


Figure 14.3 TAI-UT1 and TAI-UTC.

14.4

UT1

With the introduction of the new definition of UTC, a new variable, DUT1, was introduced as the expected difference between UT1 and UTC accurate at the level of 0.1 s. DUT1 was designed to provide a low-accuracy estimate of UT1 primarily for celestial navigators. This is distinguished from ΔT , which is the observed difference between UT1 and UTC. DUT1 is made available by the IERS and changes in its value are announced as required. Based on astronomical observations, the IERS decides upon the value of DUT1 and its date of introduction, and circulates the information one month in advance. In exceptional cases of a sudden change in the rate of rotation of the Earth, the IERS may issue a correction not later than two weeks in advance of the date of introduction. It is coded into some radio time signals and can be considered as a very coarse correction to UTC to give an approximation to UT1. The magnitude of DUT1 should not exceed 0.8 s, the difference between UTC + DUT1 and UT1 should not exceed 0.1 s, and the difference between UTC and UT1 should not exceed 0.9 s. A code was specified for the transmission of DUT1 with the radio time signals (Recommendation ITU-R TF.460).

14.5

UTC Worldwide

The availability of UTC as an international time scale approximating Greenwich mean solar time, with the precision of the SI second and matching UT1 to within

one second, led to its increasing acceptance by countries around the world. Over the years UTC has become either the basis for legal time of many countries, or accepted as the *de facto* basis for standard civil time. In 2007 UTC became the official time of the United States, replacing mean solar time. UTC is now generally accepted internationally as the worldwide standard for time. The use of time zones, with offsets mostly of hour increments, provides the civil time based on UTC for all the zones around the world.

14.6

Time Distribution

The distribution of time via radio time signals has been replaced largely by other means of time distribution depending on the user's accuracy requirements. The Global Positioning System (GPS) and the internet are now commonly used to obtain UTC. Chapter 16 describes these along with other methods of time transfer.

14.7

The Future of UTC – Leap Seconds or Not?

The ITU-R and the IAU continue to discuss possible changes in the definition of UTC. These discussions are largely concerned with possibly eliminating the leap seconds. As the Earth's rotational speed slows, leap seconds will be required more frequently. If the rate of change of the length of the day is on average about 1.7 milliseconds per century, by the end of the twentyfirst century the length of the day would be about 86 400.004 SI seconds, which would require inserting a leap second every 250 days. So in the middle of the twentysecond century two leap seconds would be required every year. In the twentyfifth century four leap seconds would be required every year. In two thousand years having leap seconds only at the ends of the months would not be adequate. However, there have only been two leap seconds in the period from 1999 through 2008, so during that time the Earth's rotational speed has not been slowing at the rate that would be expected from historical observations.

Communication and navigation systems exist that use time scales operationally that are independent of step changes, such as leap seconds. There are also a number of systems based on the current definition of UTC with its limitation of the difference between UT1 and UTC being less than one second. The costs of introducing leap seconds into the world timing infrastructure versus the costs of changing the system that has been in existence over 30 years are not well determined. So, at this time the discussion continues concerning the question of changing the definition of UTC and the procedure and date of such a possible change.

References

- Bartky, I.R. (2000) *Selling the True Time, Nineteenth-Century Timekeeping in America*, Stanford University Press, Stanford, p. 310.
- Blaise, C. (2002) *Time Lord, Sir Stanford Fleming and the Creation of Standard Time*, Random House, New York, p. 253.
- Fleming, S. (1878) *Uniform Non-Local Time (Terrestrial Time)*, S. Fleming, Ottawa.
- Fleming, S. (1879a) Time-reckoning. *Proc. Can. Inst.*, 1, 97–137.
- Fleming, S. (1879b) Longitude and time-reckoning. *Proc. Can. Inst.*, 1, 138–49.
- Hirsch, A. and Von Oppolzer, T. (1883) *Unification des Longitudes par l'Adoption d'un Méridien initial unique et Introduction d'une Heure Universelle*, Bureau central de l'Association géodésique internationale.
- Howse, D. (1997) *Greenwich Time and the Longitude*, Philip Wilson Publishers, London, p. 199.
- International Conference held at Washington for the purpose of fixing a Prime Meridian and a Universal Day, October 1884 – Protocols of the Proceedings.
- Nelson, R.A., McCarthy, D.D., Malys, S., Levine, J., Guinot, B., Fliegel, H.F., Beard, R.L. and Bartholomew, T.R. (2001) The leap second: its history and possible future. *Metrologia*, 38, 509–29.
- Trans. Int. Astron. Union (1968) *Proceeding of 13th General Assembly, Prague, 1967*, Vol. XIII B (ed. L. Perek), Reidel, Dordrecht, p. 181.
- Wheeler, G.M. (1885) *Report Upon the Third International Geographic Congress and Exhibition at Venice, Italy, 1881*, Government Printing Office, Washington.

15

Time in the Solar System

15.1

The Solar System

Our solar system is composed of four terrestrial planets, four giant planets, minor planets mostly between Mars and Jupiter, Kuiper belt objects beyond Neptune, including Pluto, which can be called a planet or not, according to your taste, the Oort comet cloud, and satellites around all of the planets except Mercury and Venus, and around Pluto, minor planets, and Kuiper belt objects, comets, and dust. Our solar system is part of the Milky Way galaxy and in motion around its center. Our galaxy is a small part of the Universe, which is composed of billions of stars, galaxies, star clusters, pulsars, quasars, nebulae, and other objects. It is approximately 13.5 billion years old and expanding. The distances are large and measured in large numbers of light years, the distance light travels in a year.

While knowledge of the solar system and the universe is continually improving, there is also much yet to be understood. Even in our solar system new objects are being discovered. The masses and compositions of the various bodies are not all well known. The interactive forces are not all well determined. The kinematics of the Earth is being determined with increased accuracy and complexity.

The observations that contribute to improved understanding of the solar system are made in different wavelengths of the electromagnetic spectrum from the Earth's surface, from satellites in orbit around the Earth, and from space probes traveling through the solar system. It is frequently essential to relate the timing of these observations in a space-time coordinate system that is fully relativistic and consistent with the gravitational potential and motion of the observational platform. As we consider time in the four-dimensional relativistic reference system, we must be aware of the presence of objects of large mass, their motions, and their effects on electromagnetic waves, and the motions of the objects of interest.

15.2

Pursuit of Uniformity

The calculations of positions and motions of solar system objects are based on the independent variable, time. It is assumed that that time is a uniform time in that the unit of time measure is the same at all times. A measure of the lack of uniformity of time, denoted by u , is provided by the variations of rate over specified intervals. It is also assumed that the theoretically defined time can be realized operationally by some physical source that can be used in recording the observations. To determine improved ephemerides the observations of bodies must be compared to predicted positions at the same time, and the differences analyzed.

The uniformity of the time scale has been challenged repeatedly over the years. There has been a progression of different times scale used for ephemerides. As accuracies have improved there has been a growing recognition of the need for improved time scales and an increase in the complexity of the definitions and relationships. We have progressed from time solely dependent on the rotation and orbital motion of the Earth to time scales based on atomic physics and the principles of the theory of general relativity.

Table 15.1 shows the changes in the uniformity of the time scales used to compute ephemerides of solar system bodies from Ptolemy in 150 AD to the present time. The values of u for the different time scales, except for ET and TCB, are the realized values for the scales at those times when they were used in the calculation of ephemerides. For ET and TCB the values are those of the definitions. The comparison of the accuracies for Ephemeris Time and Atomic Time in 1956 indicates a reason for the short lifetime of Ephemeris Time and the emergence of the time scales based on atomic physics.

Table 15.1 Uniformity of time.

Apparent Solar Time	10^{-4}	10 s day^{-1}
Mean Solar Time	10^{-8}	0.001 s day^{-1}
Ephemeris Time	10^{-10}	$0.00001 \text{ s day}^{-1}$
Atomic Time 1956	10^{-12}	$0.0000001 \text{ s day}^{-1}$
Atomic Time 2007	10^{-15}	$0.000000001 \text{ s day}^{-1}$
TCB	10^{-30}	$0.0000000000000000000000000000001 \text{ s day}^{-1}$

15.3

Pursuit of Accuracy

In parallel with the pursuit of a uniform time scale, the quest for accuracy, not only in the second of time but also in the positions and motions of the solar system bodies, is necessary for improving the understanding of our solar system. The developments in time accuracy have been described in other chapters. The pro-

gression in the accuracy of observations and ephemerides will only be briefly outlined here and can be found in more detail in other sources (*Explanatory Supplement*, 1992). The observational techniques have ranged from the unaided eye, through the telescope, photography, charge coupled devices (CCDs), radar and laser ranging, to spacecraft visits. The accuracies have gone from arcminutes to milliarcseconds, from thousands of kilometers to millimeters.

The mathematical bases of ephemerides have progressed from epicycles to Kepler's laws, from Newton's universal law of gravity to Einstein's theory of relativity, from 'computers' that were people with pencil, paper and logarithmic tables, to calculators, to punched card equipment, to modern high speed computers. Here, the accuracies have gone from arcminutes to microarcseconds, from thousands of kilometers to millimeters.

The continuing advances in the accuracy of the observational techniques, mathematical modeling, and computational capability indeed drive the need for improved accuracy in the ephemerides.

15.4 Time and Phenomena

A variety of solar system phenomena depend on the positions of the bodies and their motions. In most cases observations of these phenomena and their timings depend on the observer's location. In many cases useful observations require knowledge of the Earth's orientation in space, precise *a priori* ephemerides of the solar system, and an accurate, uniform clock time.

15.4.1 Eclipses, Occultations, Transits

Three types of phenomena are of particular importance. An eclipse takes place when one body passes into the shadow of another body. An occultation takes place when a large body passes in front of a smaller body. Transits take place when a smaller body passes in front of a larger body.

A solar eclipse takes place when the Moon is between the Sun and the Earth and blocks the sunlight from reaching a specific location on the Earth, so it is really a kind of occultation (Figure 15.1). Because of the rapid motion of the Moon's shadow, the locations of solar eclipses are in restricted paths and very sensitive to the Earth's orientation at the time of the eclipse. The types and the geometry of eclipses are illustrated in Figure 15.2. Eclipses were a very early means of determining that the Earth was a sphere.

Lunar eclipses take place when the Earth comes between the Sun and the Moon. This means the Moon is in the Earth's shadow and wherever on Earth the Moon is above the horizon, the lunar eclipse can be observed. Thus, lunar eclipses are sensitive to the ephemerides of the Sun, Moon, and Earth, but not to the Earth orientation.

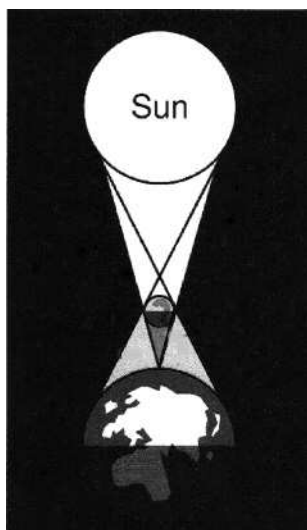


Figure 15.1 Solar eclipse.

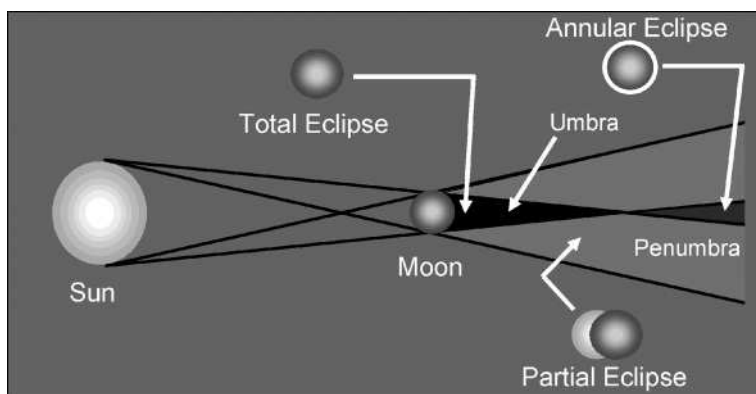


Figure 15.2 Geometry and types of solar eclipses.

Lunar occultations take place when the Moon passes in front of another celestial body. These phenomena are very sensitive to Earth orientation, star catalog positions, ephemerides of the occulted body, lunar limb profiles, and time. Occultations can provide improved knowledge concerning all these quantities. They can also provide information about the occulted body, whether it is a double star, has an atmosphere or satellite, etc. An occultation of Jupiter by the Moon is shown in Figure 15.3.

There can be transits of Mercury and Venus in front of the Sun and of satellites of Jupiter and Saturn in front of those bodies. Transits of extrasolar planets in front of stars are currently the means of detecting the existence of the smallest extrasolar planets. The timings of transits can be used to determine relative dis-

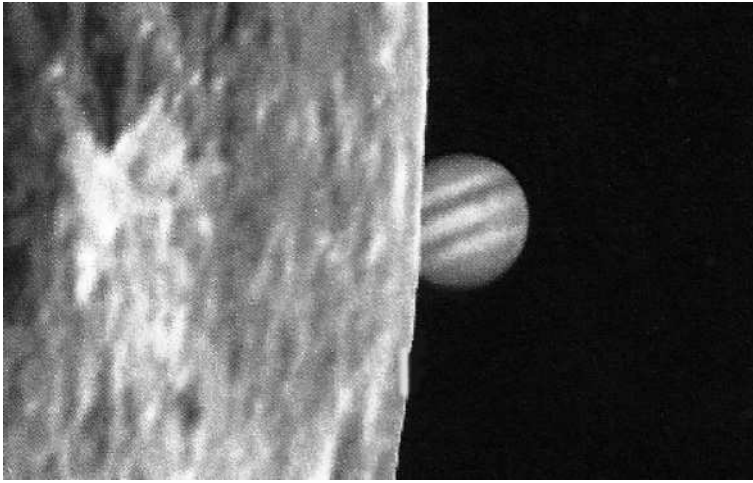


Figure 15.3 Occultation of Jupiter by the Moon.

tances. Thus, in the past, observations of the transit of Venus were made from many locations around the world in an attempt to determine an accurate value for the distance between the Earth and the Sun, also called the astronomical unit. Transits of Jupiter's satellites were used to try to determine the speed of light.

Solar eclipses are the most sensitive to variations in Earth rotation, and, as has been seen, the locations of observers of historic solar eclipses are the basis for the determination of the long-term rate of rotation of the Earth. They also provide a means of determining the long-term parameters of the ephemerides of the Earth and Moon.

15.4.2

Sunrises and Sunsets

The daily phenomena of sunrises and sunsets are the most visible evidence of the rotation and orbital motion of the Earth, as the times and locations of the phenomena change during the year. From sundials the difference between apparent and mean solar time, or the equation of time, can be detected. The atmospheric refraction effect causes differences of up to minutes in the real times of rises and sets (Figure 15.4). Hence, while the times are dependent on the Earth's orientation, the variations in Earth rotation do not have a significant effect on the times.

15.4.3

Moonrises and Moonsets

The rapid motion of the Moon is most obvious from its monthly cycle through its phases. Observations of the crescent moon immediately after new Moon are often the bases for lunar calendars, and these are very dependent on the separations of



Figure 15.4 Sunset through clouds.

the Sun and Moon and the rotation of the Earth. It is now possible to predict accurately the times for sighting the crescent Moon. Moon rises and sets are also predictable and dependent on ephemerides and Earth orientation, but the accuracies are such that variations in Earth orientation do not have a significant effect on the times.

15.5

Time and Distance

Historically the units of time, length, and mass were independent, with separate standards for each. However, atomic clocks have provided a measure of the duration of the second that is significantly more accurate than the standards for length or mass. In 1975, the 15th Conférence Générale des Poids et Mesures adopted the standard value for the speed of light based on physical measurements (SI Units) (Comptes Rendus de la 15e CGPM, 1975):

The 15th Conférence Générale des Poids et Mesures,

considering the excellent agreement among the results of wavelength measurements on the radiations of lasers locked on a molecular absorption line in the visible or infrared region, with an uncertainty estimated at $\pm 4 \times 10^{-9}$, which corresponds to the uncertainty of the realization of the metre,

considering also the concordant measurements of the frequencies of several of these radiations,

recommends the use of the resulting value for the speed of propagation of electromagnetic waves in vacuum
 $c = 299\,792\,458$ metres per second.

The combination of the knowledge of the speed of light and the standard for the second meant that the meter was better determined from those values than from any standard bar of any material in any environmental chamber. In addition, the use of radar and laser signals provided the means to measure distances from observatories to satellites, the Moon, and the planets with new precision. These measures were in terms of the round trip time for the signal up to the body and reflected down, to be observed as a return signal. Consequently, the solar system distance measurements are known more precisely than ever before using time measures and our knowledge of the speed of light.

15.5.1

Meter Definition

The meter, the fundamental unit of length in the International System of units (SI), is now defined by the speed of light in a vacuum and the SI second. It was originally defined by a prototype platinum-iridium object meant to represent $\frac{1}{10\,000\,000}$ of the length of the quarter meridian between the poles and the equator. As from 1983 the meter is defined as the length of the path traveled by light in vacuum during a time interval of $1/299\,792\,458$ of a second (SI Units).

In practice, the meter can be determined from the simple equation

$$\lambda = c_0 / f, \quad (15.1)$$

where λ is the wavelength corresponding to the frequency f of an energy level transition in a particular atom or molecule in free space, and c_0 is the speed of light in free space $= 299\,792\,458\text{ m s}^{-1}$. If the frequency f is taken as that of the hyperfine structure of cesium 133, namely $f_{\text{Cs}} = 9\,192\,631\,770\text{ Hz}$, then λ is fixed for this transition at $\lambda_{\text{Cs}} = 32.612\,255\,717\,494\,1\text{ mm}$, effectively defining the meter.

15.5.2

Radar Ranging

Radar ranging is the timing of the round trip of a radar signal from the Earth to a planet (uplink) and back from the planet to the Earth (downlink). The timing must be done by an accurate clock. To make an accurate determination of distance, analysis of the ranging time must account for the fact that the Earth

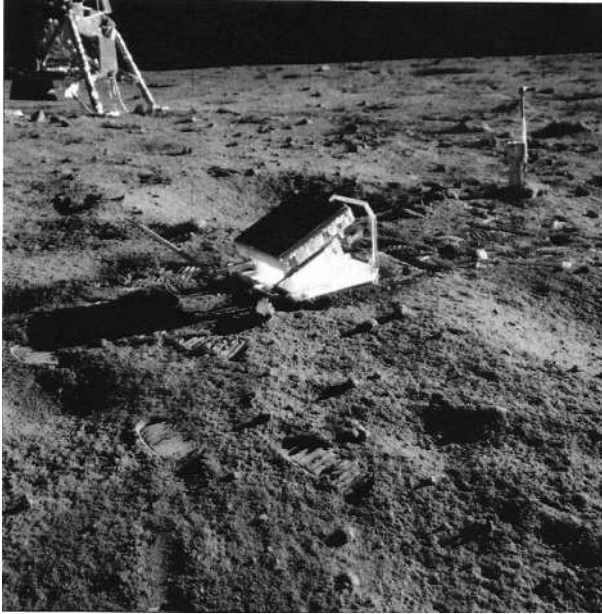


Figure 15.5 Retroreflectors placed on the Moon during the Apollo missions.

and the planet are moving and rotating during the process (see *Explanatory Supplement*, 1992).

15.5.3

Laser Ranging

Similarly to radar ranging, laser ranging is the actual timing of the round trip light signal between the Earth and the Moon or an artificial satellite. Since in this case light is being used, the equipment is very different. Telescopes are used for sending and receiving the signals and retroreflectors are required to provide a return signal. Retroreflectors were placed on the Moon during the Apollo missions (Figure 15.5) and on a Russian lunar lander. The McDonald Laser Ranging Station of the University of Texas, which was one of the first continuing lunar laser ranging operations, is shown in Figure 15.6. Since the wavelengths are much shorter than those of the radar signals, better accuracies are possible. Currently lunar laser ranging accuracies of centimeters can be achieved.

15.5.4

Navigation Systems

Determining the length of time it takes an electromagnetic signal to travel between locations and assuming a standard value for the speed of light allows us to deter-

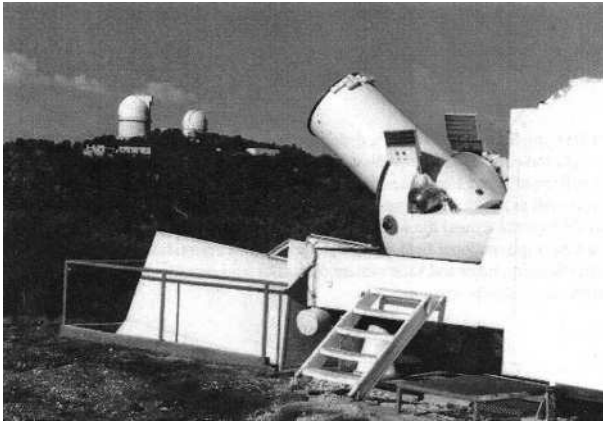


Figure 15.6 McDonald Laser Ranging Station of the University of Texas.

mine the linear distance between the two sites. Knowledge of the distances to multiple reference sites then allows us to determine a position within the reference frame of the known locations. An early application of this principle occurred during World War II with the development of the British GEE system and the American Loran system. These systems are called hyperbolic navigation systems because they make use of signals from two time-synchronized stations to determine the difference in the arrival times of the two signals, and the knowledge of that time difference places the observer on a particular hyperbolic curve between the two stations. The concept of operations is to use transmitters in ‘chains’ of at least three stations to permit the user to determine the intersection of two hyperbolic curves and thus find his or her location (Figure 15.7).

In the 1960s the U.S. Navy Timation system expanded the concept of using times of arrival of signals from precisely synchronized clocks with a limited artificial

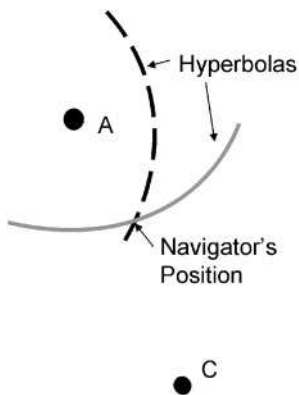


Figure 15.7 Hyperbolic navigation system. The dashed line represents the hyperbolic path defined by the difference in the time of arrival of signals from stations A and B while the grey line represents the hyperbolic path defined by the difference in the time of arrival of signals from stations A and C. Their intersection is the navigator's position.

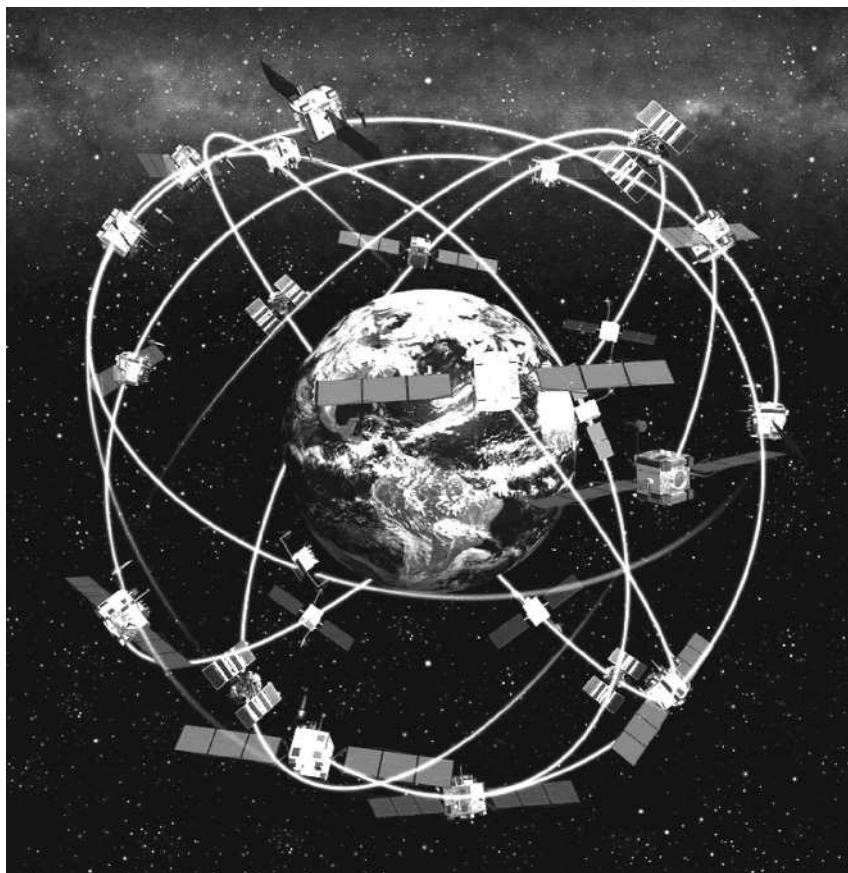


Figure 15.8 GPS constellation.

satellite system for navigation. This concept was further developed into the Global Positioning System (GPS) (Figure 15.8). Similar systems have been, or are being, developed, such as GLONASS, Galileo, and Beidou/Compass (see Section 16.3.10). The satellites each carry accurate clocks and transmit a timed signal. Receivers monitor multiple satellites and from the timed signals determine their distance from each satellite. The multiple distance measures determine the position of the receiver in three dimensions. These measurements can be used to determine the kinematics of the Earth at fixed positions on Earth as well as determining the motions of moving objects on and around the Earth.

The accuracy of the navigational solutions for global navigation satellite systems (GNSS) depends critically on the precision of the clocks whose timing signals are broadcast by the satellites. The GPS satellites currently use mostly rubidium clocks, which are used to produce an internal time scale called GPS System Time or simply GPS Time. For the clocks on board, the satellites broadcast derived corrections that are used by the GPS receivers to determine a geodetic position. The

U.S. Naval Observatory (USNO) monitors the satellite clocks to compare with its master clock, and this information is used to steer GPS Time close to UTC modulo integral seconds. (GPS Time does not insert leap seconds to maintain a close relationship to UT1. Consequently the difference between UTC and GPS Time continues to grow.)

For precision navigational systems operating in space on artificial satellites, three major systematic relativistic effects also need to be considered with regard to using the broadcast timing signal. Time dilation refers to the effect of the clock's velocity on its frequency. The effect of time dilation is given by

$$\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad (15.2)$$

where Δt is the time interval between events for an observer in some inertial frame, $\Delta t'$ is the time interval between those same events, as measured by another observer, moving with velocity v with respect to the former observer, and c is the speed of light. GPS satellites move with a velocity of 3.874 km s^{-1} , causing the clocks to appear to run slow in comparison to clocks on the Earth's surface by $7 \mu\text{s day}^{-1}$.

Because the clocks, at an altitude of 20,184 km, experience a gravity field less than clocks on Earth, the gravitational red shift causes them to run fast. This effect can be estimated roughly, assuming circular orbits for the satellites, from

$$\Delta t' = -\Delta t \left(\frac{\Delta \Phi}{c^2} \right) = -\Delta t \frac{1}{c^2} \left[\frac{-GM_E}{r} - \left(\frac{-GM_E}{a_E} \right) \right], \quad (15.3)$$

where Δt is the time interval between events for an observer in some gravitational potential, $\Delta t'$ is the time interval between those same events, as measured by another observer experiencing a difference in the gravitational potential of $\Delta \Phi$ with respect to the former observer, G is the gravitational constant, M_E is the mass of the Earth, r is the distance of the satellite from the Earth's center of mass, and a_E is the radius of the Earth's geoid. This amounts to $45 \mu\text{s day}^{-1}$ for GPS satellites. The net effect is that a GPS satellite clock runs fast by approximately $38 \mu\text{s day}^{-1}$ compared to a clock at rest on the Earth's geoid. Considering that the timing signal travels at a rate of approximately 300 m in one μs , this large effect is compensated in the satellite clocks by offsetting the frequency before launch, so that the satellite clock appears to run at the same rate as a clock on the ground. Since the satellite orbits are not truly circular, but have an orbital eccentricity of about 0.02, there is also a sinusoidal semi-diurnal variation in the clocks' timing with an amplitude of 46 ns. This correction must be calculated and taken into account in the user's receiver.

The third effect is the Sagnac delay, which is caused by the motion of the receiver on the surface of the Earth due to the Earth's rotation during the time when the signal is on its way from the satellite. This delay is computed in the GPS receiver and is given by

$$\Delta t_{\text{Sagnac}} = \frac{2\omega A}{c^2}, \quad (15.4)$$

where ω is the rotational speed of the Earth and A is the area swept out by the position vector with respect to the center of the Earth projected onto the equatorial plane. For a stationary receiver on the geoid, the Sagnac correction can be as large as 133 ns.

15.6

Space Mission Times

When spacecraft orbit around the Earth, the coordinates and times can be provided in terms of geocentric nonrotating coordinates or in terms of terrestrial coordinates that are rotating. When the spacecraft travel to other planets, barycentric coordinates are the most convenient. The observations typically are timed using UTC, and appropriate transformations can then be made. The round trip times to and from the spacecraft can be used as measures of the spacecraft position. Also the frequency of the received signal will be changed by the Doppler Effect, so the velocity of the spacecraft can be observed.

15.6.1

Doppler Effect

The Doppler Effect, which was proposed by Christian Doppler in 1842, is the change in frequency, or wavelength, of a signal as detected by an observer moving relative to the source of the signal. The total Doppler Effect results from the motion of the source and the motion of the observer. The change in frequency, Δf , is

$$\Delta f = \frac{fv}{c} = \frac{v}{\lambda}, \quad (15.5)$$

and the observed frequency is

$$f' = f + \Delta f, \quad (15.6)$$

where f is the transmitted frequency, v is the velocity of the transmitter relative to the receiver in meters per second, positive when moving toward one another, negative when moving away, c is the speed of the wave traveling in air or a vacuum, and λ is the wavelength of the transmitted wave.

15.7

Proper Times at Planets

Chapter 8 discusses the differences between terrestrial proper and coordinate time scales in comparison to Barycentric Coordinate Time TCB. Similar relationships

Table 15.2 Magnitude of terms for relativistic time scales for the Earth and Mars.

Secular drift		Maximum Amplitude	
		Diurnal term	Principal periodic term
TT-TCG	$60.2\mu\text{s day}^{-1}$	$2.1\mu\text{s}$	
MT-TCA	$12.1\mu\text{s day}^{-1}$	$0.9\mu\text{s}$	
TCB-TCG	1.28ms day^{-1}		1.7ms
TCB-TCA	0.84ms day^{-1}		11.4ms

will hold for all of the other solar system bodies. Until recently there has been no need to define such time systems for those bodies. However, the accuracy requirements for spacecraft at Mars are approaching the point where Mars Time may need to be defined and implemented in the near future.

In a manner similar to Terrestrial Time it would be possible to define Mars Time (MT), and just as we have defined Geocentric Coordinate Time (TCG), it would be possible to define Areocentric Coordinate Time (TCA) based on the Mars gravitational potential and orbital parameters. Such time scales would be used for spacecraft orbiting or vehicles on the surface of Mars. Table 15.2 indicates the size of the terms in comparison with the analogous time scales related to the Earth (Nelson, 2006). If such a time system were needed, IAU commission 52 on Relativity for Fundamental Astronomy would be likely to define appropriate Martian time scales officially. For the Moon we expect the secular drift rate between TT and time on the lunar surface to be $56.0\mu\text{s d}^{-1}$ and the amplitude of the periodic effect to be $0.48\mu\text{s}$ at the Moon's orbital period of 27.3 days (Nelson, 2006).

15.8

Pulsars – An Independent Source of Time?

Currently the assumption is that time systems based on atomic physics and those based on the dynamics of solar system bodies do not differ in a nonlinear manner. So Terrestrial Time and atomic time are related by a linear expression. Interest continues in the possibility of finding an additional independent source of time. One suggested source of such a time system is pulsars.

Pulsars are strongly magnetic rotating neutron stars that emit radiation in a beam that is only observed on Earth when that beam is pointed in the right direction. The beam rotates with the star with periods ranging from the order of milliseconds to seconds. Timings of their rotation periods are analyzed to determine possible improvements in the ephemerides of solar system bodies, and to establish evidence of gravitational waves. In addition, pulsar timings, particularly those of pulsars with periods of revolution of milliseconds, have been suggested as possible contributions to improving the long-term accuracy of time scales. Analyses of these observations require very accurate ephemerides and physical models in

order to obtain precise measures of the intervals. The ephemerides of solar system bodies and relativistic effects must be taken into account precisely. Other sources of error include receiver noise, modeling the gravity wave background, and propagation effects.

The utility of pulsar timings for constructing time scales is based on the knowledge of the total number of rotations of the millisecond pulsars. This depends on the accuracy with which the periods of their rotation can be modeled. Generally pulsars can be expected to have periods that change linearly. However, it is possible to observe apparent ‘glitches’ in the periods, so operational monitoring of pulsars is critical.

R. N. Manchester (2008) points out the fundamental differences between atomic time scales and a possible pulsar-based time scale. Such a time scale is based on the physics of massive rotating bodies and is totally isolated from the solar system or the Earth. An atomic clock has a finite lifetime on the order of a decade, but pulsars are expected to continue rotating for billions of years. Since the precision of pulsar observations is on the order of tens of nanoseconds, pulsars are currently only possibly useful to contribute stability for intervals of several years. Finally, their contributions are based on the stability of observationally determined information regarding their periods and the rate of change of these periods, and not on physical principles. Consequently, they cannot be used to check atomic timescales. However, they could possibly be used to provide long-term stability to time scales. Petit and Tavella (1996) concluded that for averaging times greater than one year, it could be possible to construct a time scale composed of the observations of an ensemble of pulsar timings. This scale could have stability better than that derived from a single pulsar and better than an atomic time scale with those averaging times. Results have been limited by the lack of observations.

References

- Comptes Rendus de la 15e CGPM (1975), 1976, 104. Available at <http://www1.bipm.org/en/convention/cgpm/resolutions.html>.
- Explanatory Supplement to the Astronomical Almanac*, 1992 (ed. P. Kenneth Seidelmann), University Science Books, Mill Valley, CA.
- Manchester, R.N. (2008) The parkes pulsar timing array project, in *40 YEARS OF PULSARS: Millisecond Pulsars, Magnetars and More*. AIP Conference Proceedings, Vol. 983, pp. 584–92.
- Nelson, R.A. (2006) Relativistic transformations for time synchronization and dissemination in the solar system. International Astronomical Union XXVIth General Assembly.
- Petit, G. and Tavella, P. (1996) Pulsars and time scales. *Astron. Astrophys.*, **308**, 290–8.

16

Time and Frequency Transfer

16.1

Historical Transfer Techniques

Time and frequency transfer refers to the techniques and models used to compare clocks and/or frequency devices. Usually such comparisons are made between a conventionally accepted standard and the user's device. The available techniques can range from a low-precision visual comparison to sophisticated systems capable of the highest precision. Historically, astronomical observations of star transits were used to adjust clock timing, and clocks driving bell towers served as local time standards adequate for most users' needs. However, as the requirements for precise time and frequency grew, the needs for improved time and frequency transfer processes grew. In the nineteenth century, when navigators were some of the most demanding users of time, a time ball was often used in seaports to distribute accurate time. A ball was dropped on a highly visible pole at a pre-arranged time allowing those who could see the ball to synchronize their clocks. With the development of the railroad and the consequent need for schedule coordination, time signal distribution via telegraph became critical (Bartky, 2000). Later, in the twentieth century, wireless (radio) time signals became the principal means of time and frequency transfer.

Today a number of time and frequency transfer techniques are available depending on users' needs. The International Telecommunications Union–Radiocommunications Sector (ITU-R) provides a handbook *Selection and Use of Precise Frequency and Time Systems* (1997) with detailed information on modern time dissemination systems. Table 16.1 outlines techniques commonly used in current practice. Each is described in the following sections. Precise time and frequency transfer using any technique, however, depends not only on the precision physically possible with the various means of sending and receiving the signals, but also on thorough attention to modeling the path of the comparison signals and careful calibration of the equipment involved.

Table 16.1 Techniques used currently in time and frequency transfer.

Type	Precision		Coverage
	Time	Frequency	
Coaxial cable	1–10 ns	10^{-14} – 10^{-15}	Local
Telephone	1–10 ms	10^{-6} – 10^{-8}	Regional
Optical fiber	10–50 ps	10^{-15} – 10^{-17}	Local
	100 ns	10^{-13} – 10^{-14}	Regional
Microwave link	1–10 ns	10^{-14} – 10^{-15}	Local
Television broadcast	10 ns	10^{-12} – 10^{-13}	Local
INTERNET	1–10 ms	10^{-6} – 10^{-8}	Global
High-frequency broadcast	1–10 ms	10^{-6} – 10^{-8}	Global
Low-frequency broadcast	1 ms	10^{-10} – 10^{-11}	Regional
Low-frequency navigation	1 μ s	10^{-12}	Regional
Navigation satellite broadcast	10–500 ns	10^{-9} – 10^{-13}	Global
Navigation satellite carrier phase	0.5–1 ns	10^{-14} – 10^{-15}	Global
Communication satellite			
Two-way	0.5–1 ns	10^{-14} – 10^{-15}	Global

16.2

Time and Frequency Dissemination Modeling

Time and frequency transfer systems can use either one-way or two-way methods of exchanging information. In the former it is necessary to estimate the delay in the propagation of the time signal. Errors due to the modeling of these estimates can range from a few nanoseconds to more than a few milliseconds depending on the technique used. Two-way techniques use the nearly simultaneous exchange of time signals along the same path. In this case, various effects, including the propagation delay, can be practically eliminated in the data processing. If timing better than $\pm 1 \mu$ s is required, relativistic effects should also be considered.

16.2.1

Propagation Effects

The nature of the medium through which the information is sent has a significant influence on the accuracy of the comparisons. The attenuation of the signals as well as delays and possible phase shifts differ among the various techniques, but experimental testing of cables and physical modeling of tropospheric and ionospheric delays generally permit users to estimate the quantitative nature of the effects and make the appropriate corrections.

Electromagnetic signals travel with the speed of light, c , in a vacuum. In other media the speed is reduced, causing delays and possible shifts in the phase of the signals. In free space we also know that the signal strength falls off following a $1/r^2$ law, where r is the distance from the transmitter to the receiver. In media such as coaxial cables or optical fibers, however, the loss of signal strength can vary considerably depending on the materials used, and must be determined experimentally.

Electromagnetic waves are usually considered to follow a line-of-sight path, but the presence of the ionosphere, with its variable electron content and changing height, complicates estimates of delays in free space. Lower frequencies, in particular, propagate via a ground wave signal that can be affected by changes in ground conductivity, solar flares, geomagnetic storms, and variations in the tilts of the layers in the ionosphere (Middleton, 2001).

16.2.2

Calibration

A critical element in accurate time comparison is the calibration of signal delays through the equipment in the chain of the time/frequency transfer. Laboratory calibration of transmitter, receiver, and cable delays can provide accuracies on the order of single nanoseconds. However, environmental changes, particularly in temperature, can produce significant systematic errors, and careful attention to time variations in the calibrated delays is often necessary for applications requiring the highest accuracy.

16.2.3

Relativistic Effects

In the language of general relativity, two clocks are said to be synchronized in a coordinate system when their readings are equal at common dates in the coordinate time t of the system. A 'coordinate clock comparison' is the difference of clock readings at the same date in t . For terrestrial applications up to the orbit of geostationary satellites, two coordinate systems are often used. Both are geocentric. One rotates with the Earth and is called the Geocentric Terrestrial Reference System (GTRS). It is the system originally defined by the IAU in 1991 and improved in subsequent years. The second, called the Geocentric Celestial Reference System (GCRS) is nonrotating. The rotating system is obtained from the latter by a change of space coordinates that does not modify the coordinate time, so using either system should lead to the same result. The following formulae assume that Terrestrial Time (TT) (practically realized by $TT = TAI + 32.184 \text{ s}$) is the coordinate time t . However, using either TT or Geocentric Coordinate Time (TCG) leads to the same results if properly handled.

Comparison of clock readings requires the evaluation of the accumulated coordinate time during the time transfer, either by the physical transportation of

a clock or the transmission of a signal. A key quantity for a clock, A , generating proper time, τ_A , is the ratio, $d\tau_A/dt$, which is, in general, variable along the world line of the clock. Discussions of relativistic time comparisons appear in *Explanatory Supplement* (1992) and in Guinot (1997).

16.2.3.1 Clock Transport in a Rotating Reference Frame

When time is transferred between two points, P and Q, using a portable clock, it is convenient to use the rotating reference frame with the potential $W(\mathbf{r})$, which includes the potential of the Earth's rotational motion. The coordinate time accumulated during transport is

$$\Delta t = \int_P^Q \left[1 - \frac{\Delta W(\mathbf{r})}{c^2} + \frac{v^2}{2c^2} \right] d\tau + \frac{2\omega}{c^2} A_E, \quad (16.1)$$

where c is the speed of light, ω is the angular velocity of the Earth's rotation, v is the speed of the clock with respect to the ground, \mathbf{r} is a vector from the center of the Earth to the clock, which is moving from P to Q, and A_E is the equatorial projection of the area swept out during the time transfer by the vector \mathbf{r} as the clock moves from P to Q. $\Delta W(\mathbf{r})$ is the potential difference between the location of the clock at \mathbf{r} and the geoid in an Earth-fixed coordinate system with the convention that $\Delta W(\mathbf{r})$ is positive when the clock is above the geoid. $d\tau$ is the increment of proper time accumulated on the portable clock as measured in the rest frame of the clock, that is, the reference frame traveling with the clock. A_E is measured in an Earth-fixed coordinate system. As the area A_E is swept out, it is taken as positive when the projection of the clock's path on the equatorial plane is eastward. For a clock above the geoid by a height, h , of 20 km, the approximation of $\Delta W(\mathbf{r})$ by gh , where g is the total acceleration due to gravity (including the rotational acceleration of the Earth) evaluated at the geoid, leads to an error which may reach about 1×10^{-14} in relative frequency. For a better approximation, the potential difference $\Delta W(\mathbf{r})$ can be calculated to greater accuracy by

$$\Delta W(\mathbf{r}) = -GM_e \left(\frac{1}{r} - \frac{1}{a_e} \right) - \frac{1}{2} \omega^2 (r^2 \sin^2 \theta - a_e^2) + \frac{J_2 GM_e}{2a_e} \left[1 + \left(\frac{a_e}{r} \right)^3 [3 \cos^2 \theta - 1] \right]. \quad (16.2)$$

where a_e is the equatorial radius of the Earth, r is the magnitude of the vector \mathbf{r} , θ is the colatitude, GM_e is the product of the Earth's mass and the gravitation constant, and J_2 is the quadrupole moment coefficient of the Earth ($J_2 = +1.083 \times 10^{-3}$). See also Klioner (1992) for an evaluation of $\Delta W(\mathbf{r})/c^2$ as a function of the altitude.

16.2.3.2 Nonrotating Local Inertial Reference Frame

When time is transferred between points P and Q by means of a clock, the coordinate time elapsed during the motion of the clock is

$$\Delta t = \int_P^Q \left[1 - \frac{V(\mathbf{r}) - U_g}{c^2} + \frac{v^2}{2c^2} \right] d\tau, \quad (16.3)$$

where $V(\mathbf{r})$ is the potential at the location of the clock and v is the velocity of the clock, both as viewed from a geocentric nonrotating reference frame (in contrast to the expression for clock transport in a rotating reference frame given in Section 16.2.3.1). U_g is the potential at the geoid, including the effect of the potential of the Earth's rotational motion. Note that $V(\mathbf{r})$ does not include the effect of the Earth's rotation and should include the tide-generating potential of external bodies. This equation also applies to clocks in geostationary orbits, but should not be used beyond a distance of about 50000 km from the center of the Earth.

16.2.3.3 Electromagnetic Signals Transfer in a Rotating Reference Frame

In a geocentric, Earth-fixed, rotating frame, the coordinate time elapsed between emission and reception of an electromagnetic signal in vacuum is

$$\Delta t = \frac{1}{c} \int_P^Q \left[1 - \frac{\Delta V(\mathbf{r})}{c^2} \right] d\sigma + \frac{2\omega}{c^2} A_E, \quad (16.4)$$

where $d\sigma$ is the increment of standard length, or proper length, along the transmission path, and $\Delta V(\mathbf{r})$ is the difference between the potential at the point \mathbf{r} and that at the geoid, as viewed from an Earth-fixed coordinate system. A_E is the area circumscribed by the equatorial projection of the triangle, whose vertices are at the center of the Earth, at the point of transmission of the signal P and at the point of reception of the signal Q. The area A_E is positive when the signal path has an eastward component.

The second term in the integral amounts to about a nanosecond for a round-trip trajectory from the Earth to a geostationary satellite. In the third term, $2\omega/c^2 = 1.6227 \times 10^{-6} \text{ ns km}^{-2}$; this term can contribute hundreds of nanoseconds for practical values of A_E . The increment of proper length $d\sigma$ can be taken as the length measured using standard rigid rods at rest in the rotating system. This is equivalent to measurement of length by taking $c/2$ times the time (normalized to vacuum) of a two-way electromagnetic signal sent from P to Q along its transmission path.

16.2.3.4 Electromagnetic Signals Transfer in a Nonrotating Local Inertial Reference Frame

In a geocentric nonrotating, local inertial frame, the coordinate time elapsed between emission and reception of an electromagnetic signal is

$$\Delta t = \frac{1}{c} \int_P^Q \left[1 - \frac{V(\mathbf{r}) - U_g}{c^2} \right] d\sigma, \quad (16.5)$$

where $V(\mathbf{r})$ and U_g are defined as above, and $d\sigma$ is the increment of standard length, or proper length, along the transmission path.

16.3

Time and Frequency Dissemination Systems

Table 16.1 shows some of the more commonly used systems for time and frequency transfer. Their capabilities vary significantly, and details regarding each of them are given below.

16.3.1

Coaxial Cable

Direct time and frequency transfer using coaxial cable connections is commonly used in laboratory settings where distances between clocks are less than a few hundred meters. The physical parameters of the cable can vary with environmental parameters, so careful attention should be given to the environmental temperature and its stability as well as the type and length of the cable when making comparisons. Extreme accuracy may require frequent calibration. Speeds of signal transmission in such cables vary, but typical values can range from 65 to 85% of the speed of light in a vacuum.

16.3.2

Telephone

Time transfer can be provided by regional telephone services. These range from a simple voice recording of the current time to coded information for use with automated equipment. Without compensation for path delays, errors can be of the order of 0.1 to 1 second. With compensation, these errors can be of the order of 1 to 10ms.

16.3.3

Optical Fiber

Optical fiber connections between clocks can provide very high accuracy capabilities for time and frequency transfer. Sub-nanosecond time transfer can be possible over relatively short distances, but careful attention to the environmental stability of the fiber is essential. It is also important that such systems be calibrated frequently to make operational use of the capability of optical fiber.

16.3.4

Microwave Links

Within local areas, microwave links can be used to distribute time and frequency with high accuracy. Two-way connections are necessary to provide the highest accuracy. Capabilities depend on atmospheric conditions as well as possible

problems with signal reflections caused by objects along the distance traveled (multipath).

16.3.5

Television Broadcast

It is possible to take advantage of the timing of local television signals for time and frequency transfer. Common-view monitoring of a television signal can be used to provide time transfer accuracy of the order of 10 ns. Each site involved in the transfer must monitor the same synchronization pulse in the television signal. The difference between the measurements then provides the time comparison. Specialized equipment is required at each site to use this technique.

16.3.6

Internet

A number of stratum 1 servers on the Internet can also be used to provide accurate time. Network Time Protocol (NTP), the most commonly used time protocol, provides the best performance. User software runs continuously as a background task that automatically updates the computer clock, using servers at many different locations around the world.

16.3.7

High-Frequency Radio Signals

High-frequency (short-wave) radio signals continue to be an easy way of receiving timing information. Table 16.2 shows the high-frequency time signals available. The transmitted time is generally UTC, with a code or voice transmission of DUT1, so that an estimate of UT1 accurate to the nearest tenth of a second can be obtained using the relation $UT1 = UTC + DUT1$. ITU Recommendation ITU-R TF.768 contains a complete listing of HF services, including details of the content and format of the broadcasts. A comprehensive listing of stations and codes that are broadcast is also available in National Geospatial-Intelligence Agency Publication 117 (2005).

Broadcasts of voice time announcements provide accuracy of the order of a few tenths of a second, but with special equipment the accuracy of the received time signal can be on the order of a few milliseconds. It is limited by unmodeled variations in the travel time of the radio signals, particularly at higher frequencies. Reception is usually better for frequencies less than 10 MHz during nighttime hours and for the higher frequencies during daytime hours. The quality of the reception depends on tropospheric and ionospheric conditions. Uncertainty in the number of reflections of the signal off of the ionospheric reflecting layers complicates the calculation of propagation delays. Generally a single reflection can be assumed for distances less than 1600 km.

Table 16.2 High-frequency time signals.

Frequency	Call sign	Location	Broadcast times
2.5 MHz	BPM	Xian, China	Various
	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous
3.33 MHz	CHU	Ottawa, Canada	Continuous
3.81 MHz	HD2IOA	Guayaquil, Ecuador	0500–1700 UTC
4.996 MHz	RWM	Moscow, Russia	Continuous
4.998 MHz	EBC	Cadiz-San Fernando, Spain	1000–1100 UTC Monday–Friday
5 MHz	BPM	Xian, China	Various
	BSF	Taipei, Taiwan	Continuous except minutes 35–40
	HLA	Taejon, Korea	Continuous
	LOL 1	Buenos Aires, Argentina	Various
	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous
	YVTO	Caracas, Venezuela	Continuous
	CHU	Ottawa, Canada	Continuous
	RWM	Moscow, Russia	Continuous
	ATA	New Delhi, India	Continuous
10 MHz	BPM	Xian, China	Various
	LOL 1	Buenos Aires, Argentina	Various
	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous
	CHU	Ottawa, Canada	Continuous
	RWM	Moscow, Russia	Continuous
	BPM	Xian, China	Various
	BSF	Taipei, Taiwan	Continuous except minutes 35–40
	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous
14.67 MHz	CHU	Ottawa, Canada	Continuous
14.996 MHz	RWM	Moscow, Russia	Continuous
15 MHz	BPM	Xian, China	Various
	BSF	Taipei, Taiwan	Continuous except minutes 35–40
	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous
15.006 MHz	EBC	Cadiz-San Fernando, Spain	1000–1100 UTC Monday–Friday
20 MHz	WWV	Fort Collins, Colorado USA	Continuous
	WWVH	Kekaha, Hawaii USA	Continuous

16.3.8

Low-Frequency Broadcast Radio Signals

Low-frequency (long-wave) signals are also used in modern time dissemination because they are not as seriously affected by ground features or buildings. They can cover larger areas than high frequency signals and pass through many types of building walls. Signal coverage can range from a few hundred to a few thousand kilometers. Signal stability is affected by ionospheric variations, and for large distances between the transmitter and the user it is important to account for the affect of sunrise and sunset along the signal path. Time signals with accuracy of the order of a millisecond are possible, but it is also possible to use these signals to calibrate the frequency of local oscillators by continuously moni-

Table 16.3 Low-frequency time signals.

Station call name	Frequency	Power	Location
DCF77	77.5 kHz	50 kW	Mainflingen, Germany
HBG	75 kHz	20 kW	Prangins, Switzerland
JJY	40 kHz	50 kW	Fukushima Prefecture, Japan
	60 kHz	50 kW	Saga Prefecture, Japan
7MSF	60 kHz	25 kW	Rugby, United Kingdom
RBV	66.66 kHz	10 kW	Moscow, Russia
RTZ	50 kHz	10 kW	Irkutsk, Russia
WWVB	60 kHz	50 kW	Fort Collins, Colorado, USA

toring the difference in phase between the signal and the local oscillator. After accounting for possible cycle slips, phase measurements with an accuracy of the order of a few tens of microseconds are possible. Table 16.3 shows providers of low-frequency timing signals.

16.3.9

Low-Frequency Navigation Signals

The LORAN-C navigation system is a low-frequency navigation signal that has also been used as a means of time transfer. It relies on synchronized broadcasts of signals from a chain of stations separated by distances of kilometers to determine the location of the receiver. The stations emit coded signals with 100 kHz bursts with varying strengths. The signals do not contain complete timing information and are not a source of UTC. However, if a user's clock is initially set to UTC by another means, Loran-C can be used to keep the clock to within a few microseconds. Although three stations are required for navigation, only one is necessary for time and frequency transfer. Specially designed timing receivers are used to track the third sub-pulse of the received bursts. Each LORAN-C station has cesium standards that are used to synchronize the signals and these are related to national standards.

Coverage is variable, depending on radiated power and the surface conductivity along the path between transmitter and receiver. The accuracy can be of the order of 50 ns to 300 ns. The limiting source of error is modeling the variations in path delay due mainly to weather-related events.

Enhancements to the Loran system are being implemented so that broadcasting stations, each having three cesium standards, will be provided with improved transmitters and measurement systems. Stations are synchronized using signals from satellite navigation systems and will provide a new ninth pulse in the broadcast signal that can be used for precise timing. A broadcast message provides information on leap seconds, differential corrections, and station identification. In the best cases it is expected that accuracies of the order of tens of nanoseconds

will be possible depending on the distances from the transmitters (<http://www.navcen.uscg.gov/eLoran/overview.htm>).

16.3.10

Navigation Satellite Broadcast Signals

Navigation satellite systems use atomic standards on board the component satellites to provide global navigation solutions. This timing capability also provides the possibility of cheap, global, highly accurate time and frequency transfer. Reception of signals from four satellites is required to determine the three-dimensional position of the user's receiver and the user's time. Timing information can be obtained from a single satellite if the user's position is already known. Results can be affected by spurious reflections of the transmitted signals, but generally timing accuracies of 20–500 ns and frequency accuracies of 10^{-9} to 10^{-13} are possible. The accuracy of timing data obtained directly depends on the quality of the receiver.

16.3.10.1 Global Positioning System

One of the most common means of obtaining accurate timing information is through the Global Positioning System (GPS) (www.gps.gov). The current GPS constellation consists of at least 28 satellites, each carrying multiple cesium and/or rubidium atomic clocks. They orbit at an altitude of approximately 20 200 km, completing two orbits per sidereal day. Their signals are broadcast in L band with frequencies ranging from 1176.45 to 1575.42 MHz. Time transfer accuracy on the order of tens of nanoseconds is possible. That figure can be improved further by using more sophisticated techniques, such as when timing laboratories compare clocks by using a common satellite clock as a means of comparison.

GPS provides accurate UTC, but the system makes use of an internal time scale called GPS System Time, or sometimes GPS Time. GPS Time is maintained close to UTC, or TAI, modulo one second, the offset being less than one microsecond and generally of order of tens of nanoseconds. The clocks of the GPS system, located in monitoring sites and in the satellites, are used to realize this internal scale. Leap seconds are not inserted in this time scale, and its starting epoch is midnight of 5/6 January 1980, so that TAI is ahead of GPS Time by 19 s, a nearly constant value. User receivers, however, generally provide UTC, making use of information that the satellites broadcast containing the time offset between UTC and GPS Time.

16.3.10.2 GLONASS

The Russian Global Navigation Satellite System (GLONASS) provides services similar to GPS. However, it uses Moscow Time (UTC + 3 h) as its time reference instead of an internal time scale. Thus, this system is directly affected by leap second insertions. General information is provided at www.glonass-center.ru.

16.3.10.3 GALILEO

The European system of satellite positioning (GALILEO), which should be operational in 2011, will provide time information similar to that of GPS.

16.3.10.4 Beidou/Compass

The Chinese navigation satellite system Beidou/Compass, which is under development, has 4 prototype geostationary satellites, the first launched on October 31, 2000. The system will consist of 5 geostationary satellites and 30 medium-height satellites, transmitting signals on the carrier frequencies: 1195.14–1219.14MHz, 1256.52–1280.52MHz, 1559.05–1563.15MHz and 1587.69–1591.79MHz. A China Satellite Navigation Project Center (CSNPC) reportedly will take charge of the research, building, and management of the system, which is supposed to be operational in 2010.

16.3.11

Navigation Satellite Carrier Phase

In addition to using the navigation satellite-transmitted codes for time and frequency transfer, it is also possible to use the phase of the carrier frequency itself to disseminate time and frequency information. Since the carrier frequency is about 1000 times higher than the frequency of the timing code, the carrier-phase methods have much higher resolution. Clocks at two separated sites observe the same satellites at the same time recording the measured phase difference between the carrier and the local frequency reference. These data can be analyzed using post-processed satellite ephemerides, ionospheric and tropospheric models to provide estimates of the time and frequency differences between the two sites. While time and frequency transfer based on the timing codes provides unambiguous measures of the time delay between the satellite and the receiver, the carrier phase measures contain unknown multiples of 2π radians that correspond to the integral wavelengths of the carrier signal between the satellite and the receiver. These are estimated in the analyses to provide high precision timing data (Ray and Senior, 2003; Delporte *et al.*, 2008).

16.3.12

Two-Way Satellite Time and Frequency Transfer

The two-way satellite time and frequency transfer (TWSTFT) technique is used to provide highly accurate time and frequency transfer independent of navigational satellite techniques. TWSTFT makes use of geostationary communications satellites to transmit simultaneously spread-spectrum timing codes using specialized modems between two users desiring to compare their clocks (see Figure 16.1). The process does not depend on knowing the satellite's position with high precision and takes advantage of the fact that the most significant propagation errors cancel. Accuracy at the sub-nanosecond level has been achieved when the locations of the antennae are well known and the electronic equipment is carefully cali-

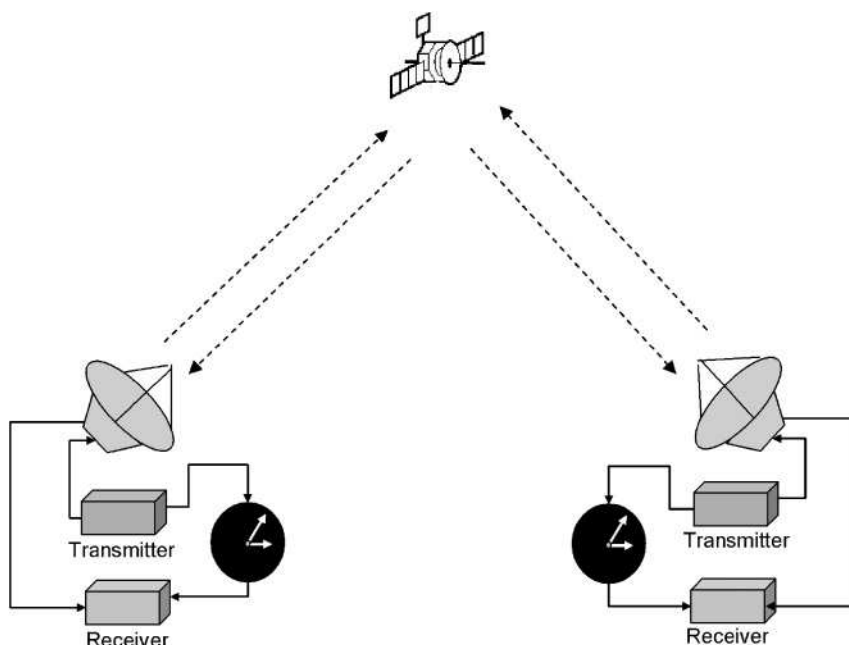


Figure 16.1 Two-way satellite time and frequency transfer.

brated. The process requires the exchange of data after the measurements have been made. Asymmetric delays in the communication satellite can be a source of error, and relativistic corrections should be applied to achieve the highest precision (Kirchner, 1999; Schaefer *et al.*, 2000).

References

- Bartky, I.R. (2000) *Selling the True Time, Nineteenth-Century Timekeeping in America*, Stanford University Press, Stanford, p. 310.
- Delporte, J., Mercier, F., Laurichesse, D. and Galy, O. (2008) GPS carrier-phase time transfer using single-difference integer ambiguity resolution. *Int. J. Navig. Obs.*, 2008, Article ID 273785, 7.
- Explanatory Supplement to the Astronomical Almanac* (1992) (ed. P.K. Seidelmann), University Science Books, Mill Valley, CA.
- Guinot, B. (1997) International report: application of general relativity to metrology. *Metrologia*, 34, 261–90.
- ITU Recommendation ITU-R TF.768. Available at <http://www.itu.int/rec/R-REC-TF.768-6-200305-1/en>
- Kirchner, D. (1999) Two-Way Satellite Time and Frequency Transfer (TWSTFT): principle, implementation and current performance, in *Review of Radio Science 1996–1999* (ed. W.R. Stone), Oxford University Press, Oxford, England.
- Klioner, S.A. (1992) The problem of clock synchronization—a relativistic approach. *Celest. Mech. Dyn. Astr.*, 53, 81–109.
- Middleton, W.M. (ed.) (2001) *Reference Data for Engineers Radio, Electronics, Computer & Communications (Reference Data for Engineers)*, 9th edn, Newnes, p. 1672.
- National Geospatial-Intelligence Agency Publication 117 (2005) Radio navigational aids. Available at <http://www.nga.mil/portal/site/maritime/>

- Ray, J. and Senior, K. (2003) IGS/BIPM pilot project: GPS carrier phase for time/frequency transfer and timescale formation. *Metrologia*, **40**, S270–88.
- Schaefer, W., Pawlitzki, A. and Kuhn, T. (2000) New trends in two-way time and frequency transfer via satellite. Proceedings of the 31st Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting: 1999, Dana Point, California, pp. 505–14.
- Selection and Use of Precise Frequency and Time Systems (1997) Available at <http://www.itu.int/publ/R-HDB-31-1997/en>

17

Modern Earth Orientation

17.1

Terrestrial to Celestial Reference Systems

The need for celestial and terrestrial reference systems is well established, and the mathematical procedures by which they can be related have become standardized. However, the different motions affecting these relationships still cannot be predicted with the accuracy needed for the most demanding applications. The details of the physical causes for the motions are not all well known. Consequently observations are required to augment the models in order to provide the highest level of accuracy possible. These are implemented in established procedures for the transformations between the different reference systems.

Chapter 5 describes the concepts of the procedures used to transform from a terrestrial reference system (TRS) to the celestial reference system (CRS) at any epoch t . There we saw that

$$[\text{CRS}(t)] = Q(t) R(t) W(t) [\text{TRS}(t)], \quad (17.1)$$

where $Q(t)$, $R(t)$ and $W(t)$ are the transformation matrices describing the motion of the precession/nutation, the rotation of the Earth around the axis of the pole, and polar motion, respectively (McCarthy and Petit, 2004), and t is defined by

$$t = [\text{TT} - 2000 \text{ January } 1, 12\text{h TT in days}] / 36525. \quad (17.2)$$

Note that 2000 January 1.5 TT = Julian Date 2451545.0 TT.

The precession/nutation rotation is given by

$$Q(t) = \begin{bmatrix} 1 - aX^2 & -aXY & X \\ -aXY & 1 - aY^2 & Y \\ -X & -Y & 1 - a(X^2 + Y^2) \end{bmatrix} \cdot \begin{bmatrix} \cos s & \sin s & 0 \\ -\sin s & \cos s & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (17.3)$$

$$a = \frac{1}{2} + \frac{1}{8}(X^2 + Y^2), \quad (17.4)$$

where X and Y are the angular ‘coordinates’ of the Conventional Intermediate Pole (CIP) in the CRS, provided in part by the conventional mathematical models, and

s is given by Eqs. (5.3) to (5.6). For the highest precision it is necessary to account for the differences between the observed value and the theoretical model. Mathematically,

$$\begin{aligned} X &= X_{\text{MODEL}} + \delta X \\ Y &= Y_{\text{MODEL}} + \delta Y \end{aligned} \quad (17.5)$$

where X_{MODEL} and Y_{MODEL} are the values provided by the models and δX and δY are the celestial pole offsets determined from astronomical observations.

The Earth's rotation angle is handled by the matrix:

$$R(t) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (17.6)$$

θ being the Earth Rotation Angle given by

$$\theta(T_u) = 2\pi \left(\text{UT1 Julian Days elapsed since } 2451545.0 + 0.7790572732640 \right. \\ \left. + 1.00273781191135448 T_u \right). \quad (17.7)$$

where $T_u = (\text{Julian UT1 date} - 2451545.0)$, and $\text{UT1} = \text{UTC} + (\text{UT1} - \text{UTC})$.

Finally, the polar motion rotation is given by

$$W(t) = \begin{bmatrix} \cos s' & -\sin s' & 0 \\ \sin s' & \cos s' & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & \sin \gamma \\ 0 & -\sin \gamma & \cos \gamma \end{bmatrix} \cdot \begin{bmatrix} \cos x & 0 & -\sin x \\ 0 & 1 & 0 \\ \sin x & 0 & \cos x \end{bmatrix}, \quad (17.8)$$

where x and γ are the angular 'coordinates' of the Conventional Intermediate Pole in the Terrestrial Reference System (TRS) and s' in microarcseconds, μas , can be approximated for the twenty-first century by

$$s' = -47 \mu\text{as } t, \quad (17.9)$$

where t is given in Eq. (17.2) and the angles X , Y , $\text{UT1} - \text{UTC}$, x and γ collectively are known as the Earth orientation parameters, which must be determined by observation.

17.2

Determination of Earth Orientation Parameters

Chapter 5 outlines some of the techniques that have been used in the past to provide the Earth orientation data. These have included visual and photographic telescopic observations, Doppler observations of artificial satellites, laser ranging to the Moon and artificial satellites, very long baseline interferometry, and analysis of the orbits of navigational satellites. Not all of these techniques continue to be used. A variety of high-precision techniques are used currently to relate the CRS and the TRS and to make predictions of future Earth orientation. These are discussed below.

17.2.1

Very Long Baseline Interferometry (VLBI)

The VLBI technique makes use of multiple radio telescopes observing very distant radio sources to determine the Earth's orientation with respect to the celestial reference frame realized by the conventionally adopted radio source positions. The baselines defined by the vectors between the positions of the telescopes, rigidly attached to the Earth's surface, are defined in the terrestrial frame (Boucher *et al.*, 2004). The radio sources are quasars, which are powerful radio emitters located at such great distances from the Earth that they can be considered to show minimal space motion. Monitoring of the changing aspect of the baseline vectors with respect to these sources provides the Earth orientation parameters required to implement the mathematical relationship shown in Eq. (17.1) (see Johnston, 1979). The concept is shown in Figure 17.1.

The radio sources are called 'quasi-stellar radio sources' or 'quasars' for short. They are thought to be the centers of some distant galaxies, probably with massive black holes at the center. Those that appear to be the most 'point-like,' with minimal observed space motion and variations in observed signal intensity, form a group that defines the International Celestial Reference Frame (ICRF) (Ma and Feissel, 1997). These objects broadcast powerful radio signals composed essentially of random noise.

The radio telescopes that are used to monitor the signals have diameters on the order of tens of meters and are located at continental distances to provide a better representation of the global Earth orientation and better resolution of the radio sources themselves. Their positions are well determined and provide part of the definition of the International Terrestrial Reference Frame. Figure 17.2 shows the

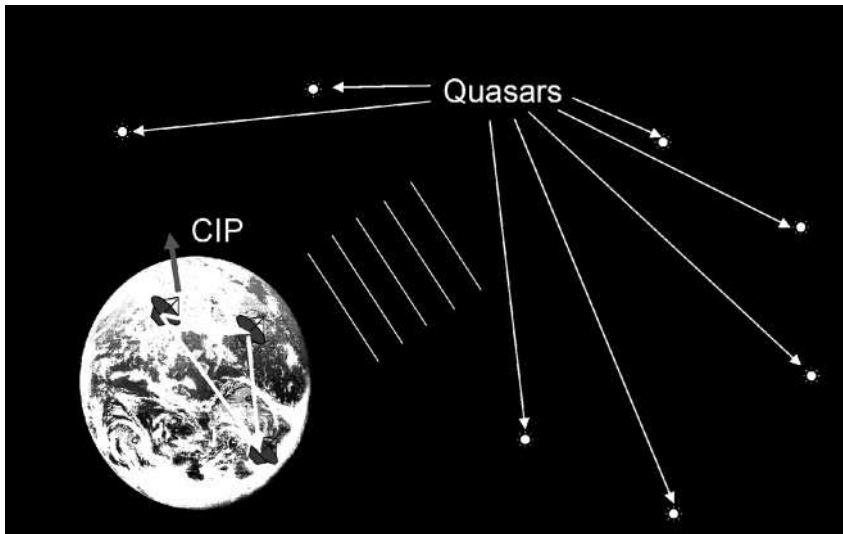


Figure 17.1 The concept of VLBI used to monitor the Earth orientation parameters.

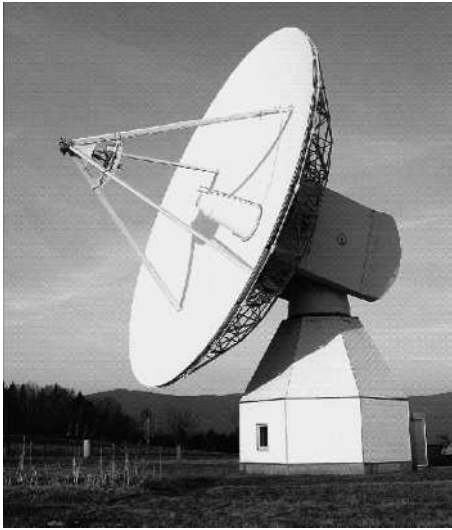


Figure 17.2 VLBI antenna at Wettzell, Germany.

telescope located at Wettzell, Germany. Figure 17.3 shows the worldwide distribution of the antennae currently in use in geodetic VLBI operations and those sites that have been used in the past for such operations. The operations are organized through the International VLBI Service for Geodesy and Astronomy (IVS) (see Chapter 18).

Observations are made by multiple telescopes observing quasars widely distributed over the sky in order to determine all of the Earth orientation parameters with the best precision. Because the telescopes are located at different distances from the radio sources there are differences in the times of arrival of the signals at each telescope. Each site has a hydrogen maser that serves as a clock to measure precisely the difference in the arrival times. The digitized signals and the times are recorded on electronic media at each telescope and then sent to a correlator, where the signals are analyzed to determine the delays and the rate of change of the delays among the telescopes involved in the observation. The development of high-speed data transfer networks permits near real-time data transfer without the necessity of recording the data on electronic media that have to be shipped to the correlator (Carter and Robertson, 1986; Carter, Robertson and Fallon, 1989).

At the correlator the data are played back. The estimated delay can be calculated using the known positions of the telescopes and quasars and *a priori* estimates of the Earth orientation parameters. This information is used to determine the actual delays and delay rates, which can then be analyzed to determine the corrections to the *a priori* estimates of the Earth's orientation. Figure 17.4 outlines the process graphically. Observations are typically made at frequencies of S band (2.3 GHz) and X band (8.4 GHz) in order to take advantage of the dispersive nature of the

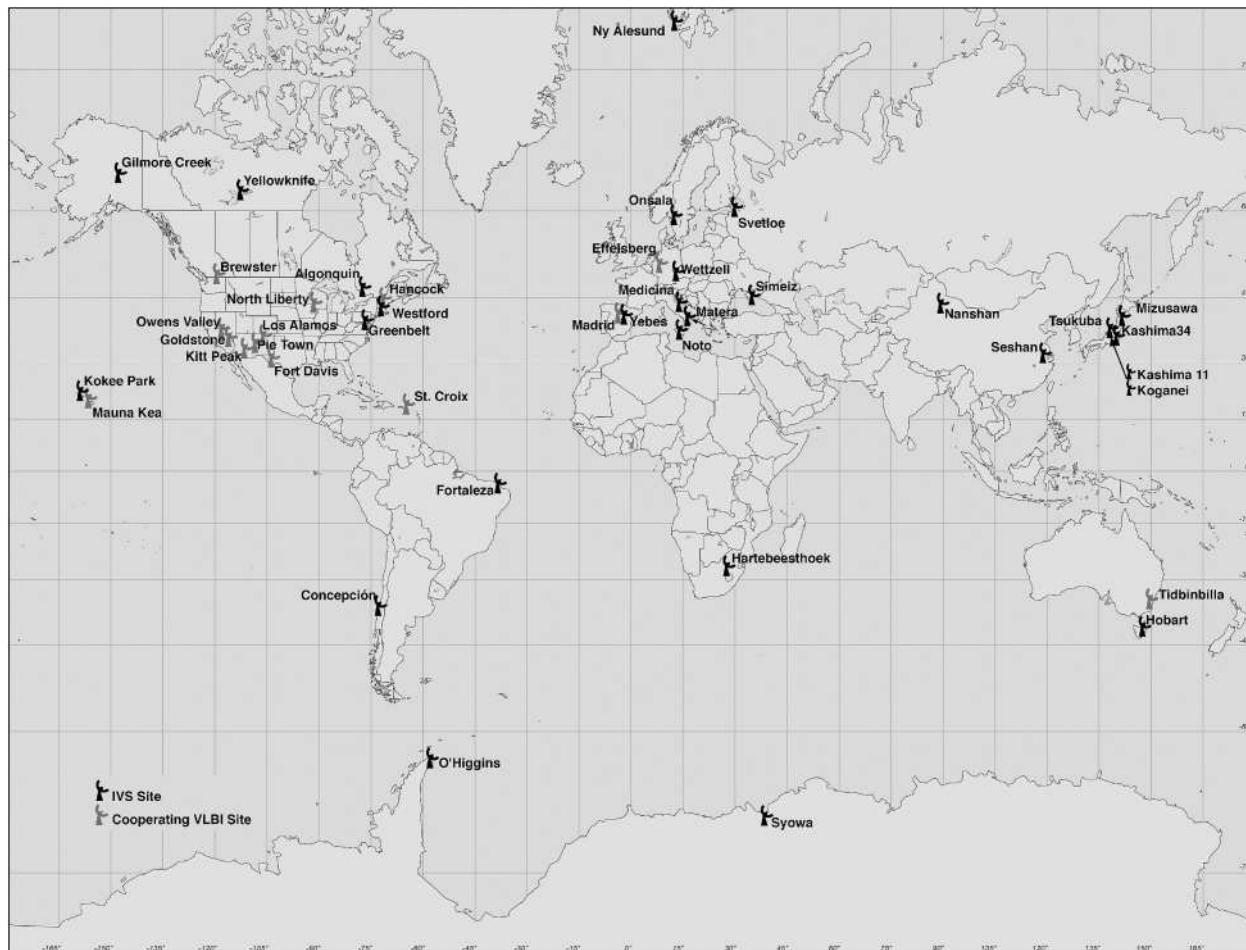


Figure 17.3 Distribution of geodetic VLBI telescopes.

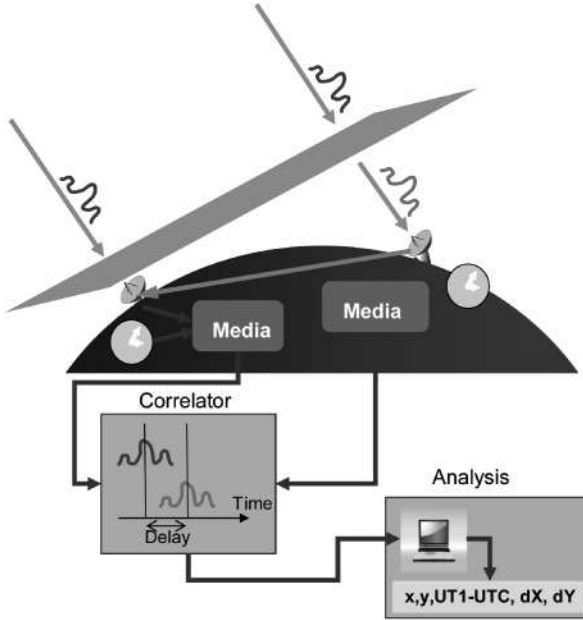


Figure 17.4 VLBI Process.

ionosphere and calculate the delays due to the total electron content (TEC) in the ionosphere. This delay, $\tau_{\text{ionosphere}}$, depends on the frequency of the signal, f , and is proportional to the TEC and inversely proportional to f^2 .

$$\tau_{\text{ionosphere}} \propto \frac{\text{TEC}}{f^2}. \quad (17.10)$$

Two general types of delay measurements are possible using VLBI: group delay and phase delay. If ϕ is the measured difference in phase (in radians) between the signals at the two telescopes, the phase delay is given by

$$\tau_p = -\frac{\phi(\omega)}{\omega}, \quad (17.11)$$

and the group delay is given by

$$\tau_g = -\frac{d\phi(\omega)}{d\omega}. \quad (17.12)$$

Phase delays can be measured with precisions of a picosecond or better, but with an unknown integral number of phase cycles, which are 451 ps at S band and 122 ps at X band. Group delays can be measured with precisions of a few picoseconds without cycle ambiguities and these are the measures typically used in geodetic VLBI observations.

VLBI precision allows analysts to determine the relative distances between the telescopes to a few millimeters and the positions of the radio sources to fractions of a milliarcsecond. Because the telescopes are fixed rigidly in the terrestrial reference frame, the variations in the observed delays provide the information necessary to determine the orientation of the baseline in the celestial frame. From these data corrections to the *a priori* estimates of x , y , UT1–UTC, X , and Y can be derived. The VLBI operations and analyses are coordinated internationally through the International VLBI Service for Geodesy and Astrometry (IVS) (see Chapter 18) (Schlüter and Behrend, 2007) (ivscc.gsfc.nasa.gov/).

Mathematically, referring to Figure 17.5, where \mathbf{B} and \mathbf{s} are the baseline vector and unit vector in the direction of the source respectively, the delay can be expressed as

$$\tau = \frac{\mathbf{B} \cdot \mathbf{s}}{c} = \frac{|\mathbf{B}| |\mathbf{s}| \cos \theta}{c}, \quad (17.13)$$

where c is the speed of light and θ is the angle between \mathbf{B} and \mathbf{s} . The station coordinates used to create \mathbf{B} must be corrected for tectonic plate motion, the tides, and the loading of the site by the atmosphere and possibly nearby oceans. \mathbf{s} is computed using the conventionally adopted precession-nutation and the *a priori* Earth orientation estimates. Observational estimates of τ must be corrected for atmospheric and ionospheric delays as well as the delays caused by the gravitational attraction of solar system bodies along the path of the signal. The observations themselves are used to calculate the atmospheric and ionospheric effects.

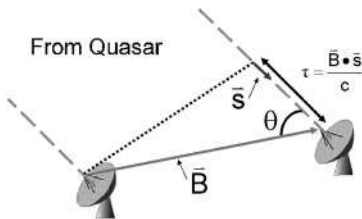


Figure 17.5 Mathematical representation of VLBI delay.

Observations can be affected by instrumental errors, propagation modeling errors, and the fact that the sources may exhibit some extended structure. The instrumental errors include possible deformation of the structure due to temperature and wind loading and possible clock errors. The delay caused by the neutral atmosphere depends on temperature, pressure, and humidity along the path of the signal, and can vary significantly with direction and time at the observing site. Sources are chosen to be used in the observations with concern for any evidence of motion or extended structure. Only the most ‘point-like’ are used to mitigate issues regarding source structure.

17.2.2

Global Positioning System (GPS)

Although VLBI is the only technique that provides all five of the Earth orientation parameters, the Global Positioning System (GPS) now provides the most accurate information regarding the polar motion portion (x , y) of the Earth orientation parameters. Monitor stations, with locations well known with respect to the terrestrial reference system, receive signals from the satellites. This information is used to determine the precise orbits of the satellite in an inertial system. As part of the process the polar motion parameters can be derived by relating the satellite positions with the locations of the monitor stations (see Figure 17.6).

The satellites orbit the Earth in planes that rotate in space with motions that are difficult to predict. The motion of the planes cannot be separated strictly from variations in the Earth's rotation. Consequently GPS cannot be used to provide accurate UT1–UTC without some form of *a priori* knowledge of that part of the rotation of the planes that is independent of the Earth's rotation. The daily rate of UT1–UTC is equivalent to the excess length of day (LOD), and this can be estimated. By integrating the LOD time series for relatively short periods of time it is possible to determine a UT1–UTC time series if an accurate integration constant can be determined using VLBI observations.

The determination of the satellite orbits requires accurate knowledge of the Earth's gravitational field including any time variations. While the orbits are sensitive to the Earth's center of mass, they are relatively insensitive to elements of the

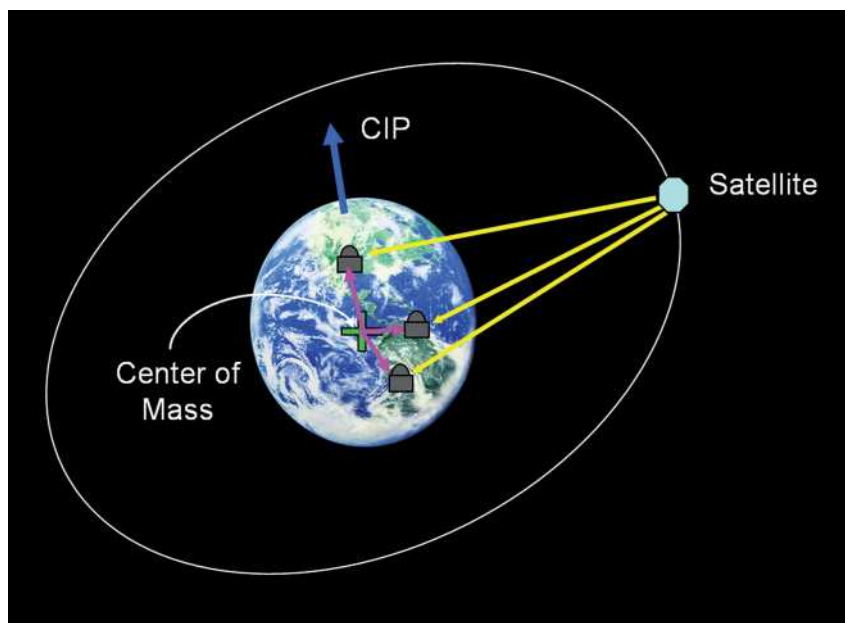


Figure 17.6 The concept of GPS used to monitor the Earth orientation parameters.

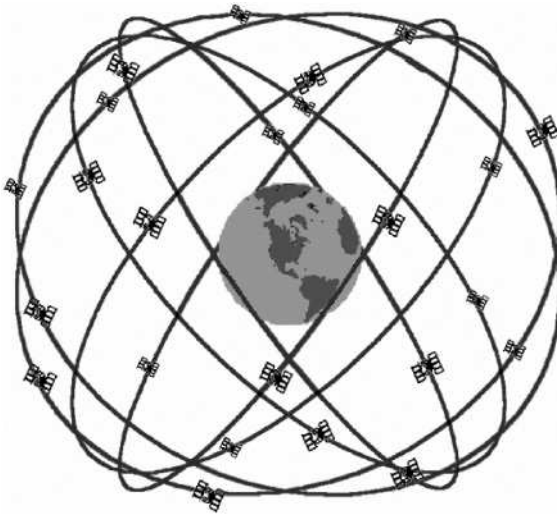


Figure 17.7 The GPS constellation of satellites.

celestial reference system. As a result the GPS orbits can provide no information regarding the celestial pole offsets, X and Y .

The constellation of GPS satellites is shown in Figure 17.7. It is composed nominally of a set of six orbital planes with inclinations of 55 degrees, each being populated with four satellites. The satellites orbit at an altitude of 20 200 km. In reality, more than 24 satellites are in orbit at any one time in order to ensure continuous operation of the system. GPS is one of a set of operational, or proposed, navigational satellite systems, but it is currently the only one that has been used operationally to provide reliable Earth orientation information.

The observations that are analyzed for Earth orientation information are coordinated by the International GNSS Service (IGS), the letters GNSS representing the Global Navigation Satellite Services (see Chapter 18). Figure 17.8 shows the network of sites that participate in this endeavor.

The accuracy of the polar motion observations is made possible by tracking the phase of the signals broadcast from the satellites. Currently the GPS satellites broadcast signals at two L-band frequencies (1.227 and 1.575 GHz). These are modulated by a pseudo-random noise code. Most of the ionospheric delay in the reception of the signals can be eliminated by using the two frequencies, just as in the case of the analysis of the VLBI observations. Tracking the phase of the carrier signal provides no information regarding the time of transmission. That information is only provided by the code. Carrier-phase measurements are differences in carrier-phase cycles and fractions of cycles over time. The carrier phase data and code information gathered from the observing sites shown in Figure 17.8 are used by the analysis centers of the IGS to derive ionospheric maps, orbits of the satellites, and Earth orientation information, as well as to maintain a precise terrestrial reference frame.

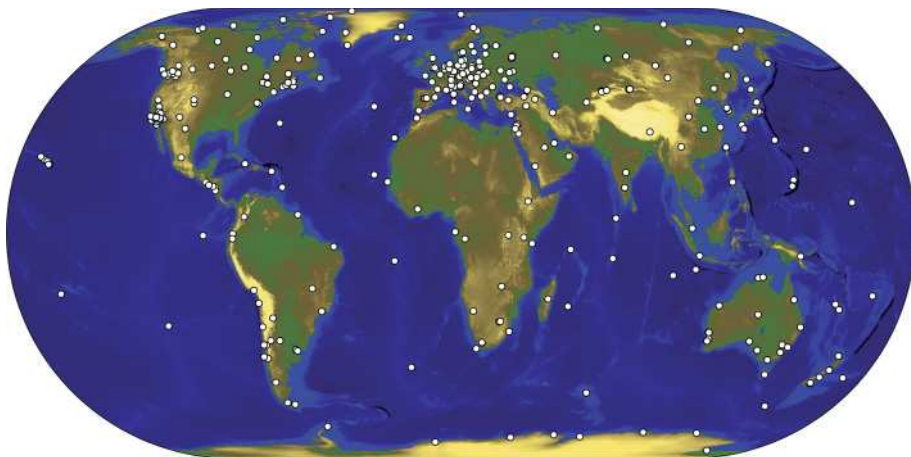


Figure 17.8 The IGS network of ground stations.

The accuracy of the results depends on the accuracy of various physical models used in the orbital analyses. These include models of the tropospheric delays, geopotential, solid Earth tides, ocean tides, ocean and atmospheric loading at the observing sites, solar pressure, atmospheric drag, relativistic effects, and Earth albedo (Rothacher, 1999; Kouba *et al.*, 2000).

17.2.3

Satellite Laser Ranging (SLR)

Satellite laser ranging (SLR) is a technique that uses a network of special-purpose telescopes to measure the travel times of very short laser light pulses to and from a set of artificial Earth satellites equipped with laser retro reflectors. Just as in the case of the GPS analyses, these data are used to relate the orbits of the satellites determined in inertial space to a terrestrial reference frame defined by the accurate locations of the telescopes that contribute their observations. The concept is shown in Figure 17.9. These observations, like the GPS observations, are only used to provide polar motion and length-of-day observations.

Typically pulses of light with wavelengths of 532 nanometers (nm) (green) and pulse lengths of 10–100 picoseconds (ps) are sent to these satellites, and the length of time for the signal to be returned to the telescope is measured with nanosecond precision. Figure 17.10 shows the telescope used to carry out these observations at the University of Texas McDonald Observatory, and Figure 17.11 shows the geographic distribution of sites where observations have been made. Not all of these sites have been used for Earth orientation determinations. Many have been used in various geodetic research efforts and to establish an accurate terrestrial reference frame (Schutz *et al.*, 1989).

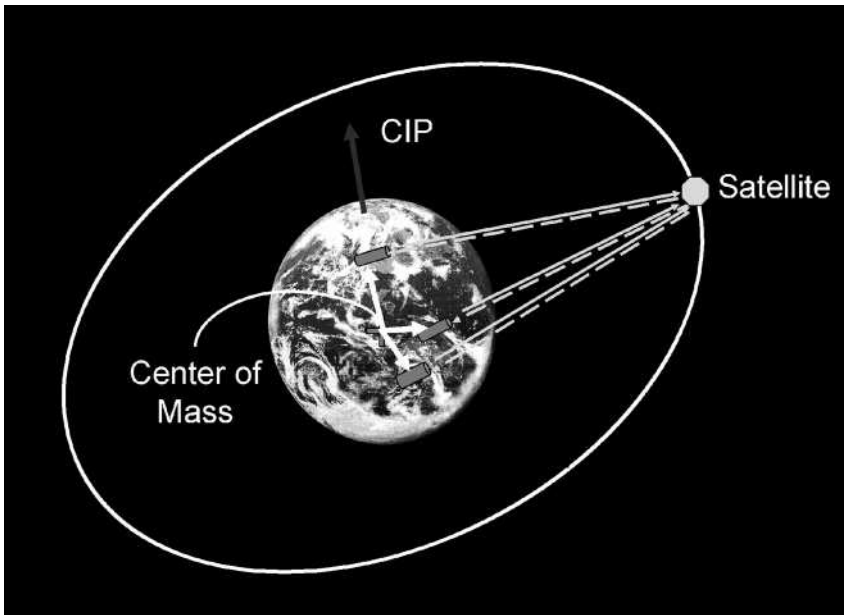


Figure 17.9 The concept of SLR used to monitor the Earth orientation parameters.

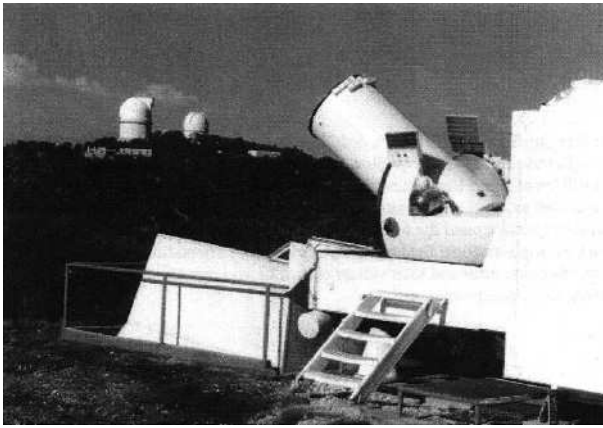


Figure 17.10 University of Texas SLR telescope at McDonald Observatory.

The satellites primarily used for Earth orientation measurements are the LAGEOS (Laser Geodynamics Satellite) satellites, which are essentially spherical satellites covered with laser retro reflectors designed to improve the signal-to-noise ratio of the observations (see Figure 17.12). LAGEOS-I was built by NASA and launched in 1976 into a nearly polar circular orbit with an altitude of 6000 km.

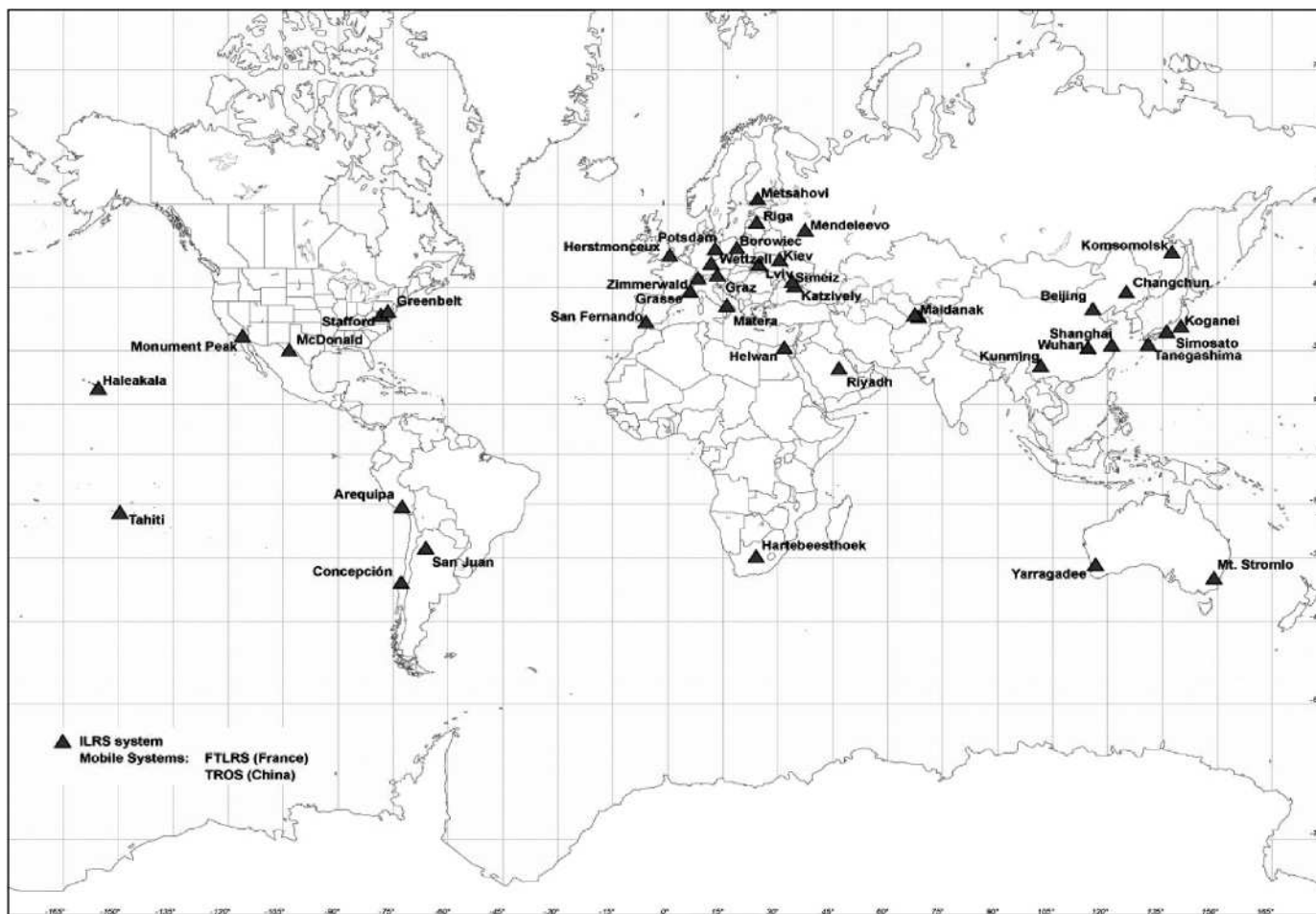


Figure 17.11 Distribution of ILRS sites.



Figure 17.12 LAGEOS satellite.

A second satellite of the same type (LAGEOS-II) was built by the Agenzia Spaziale Italiana of Italy and launched in 1992 into a similar orbit, but with an inclination of 51 degrees. Both are approximately 60 cm in diameter, weigh 411 kilograms, and are covered by 426 retro reflectors. Other satellites that have been used include Starlette (1000 km) and Stella (800 km) developed and launched by France, Etalon-I and -2 (19000 km) developed and launched by the former USSR, and Ajisai (1500 km) developed and launched by Japan.

As was the case with GPS, the Earth orientation information is derived in the process of determining the satellite orbits using the ranges from the network of observing locations. However, in the case of SLR, ranges measured using visual light are used instead of ranges measured using radio transmissions. As a result, somewhat different physical effects must be accounted for in the SLR analyses. While both techniques are sensitive to the geopotential, solid Earth tides, ocean tides, ocean and atmospheric loading at the observing sites, solar pressure, atmospheric drag, relativistic effects, and Earth albedo, the SLR technique does not have to contend with the ionospheric problems. On the other hand, it does require clear skies to receive returns.

The International Laser Ranging Service (ILRS) was created in 1998 to coordinate the observations and analyses (Gurtner *et al.*, 2005) (see Chapter 18). It provides data regarding the Earth orientation and terrestrial reference frame to the IERS routinely. These data have also been used to produce a long-wavelength gravity field reference model which supports all precision orbit determinations and provides the basis for studying temporal gravitational variations due to mass redistribution and accurate determinations of tectonic plate motions.

In the past, Earth orientation information was also provided by ranging to targets placed on the Moon's surface by the Apollo astronauts. While ranging to

lunar targets continues, operational Earth orientation information is no longer contributed from the lunar ranges.

17.2.4

Doppler Orbit Determination and Radiopositioning Integrated on Satellite (DORIS)

A fourth technique, Doppler Orbit determination and Radiopositioning Integrated on Satellite (DORIS), also contributes data to the IERS, primarily for the extension of the terrestrial reference frame. While it can also be used to determine polar motion, the precision of the derived x , y coordinates is not adequate for operational application in the area of Earth orientation parameters.

DORIS is a French system developed by the Centre National d'Etudes Spatiales (CNES) in conjunction with the Institut Géographique National (IGN) and the Groupe de Recherche de Géodésie Spatiale (GRGS). It uses specialized satellites to monitor signals from beacons placed at a number of locations worldwide. As the satellite approaches the beacon on the ground the Doppler Effect shifts the frequency seen on the satellite to a higher frequency than that actually broadcast. When the satellite moves away from the beacon, the observed frequency is lower, again as predicted by the Doppler Effect. This information, when combined with other data from a network of beacons (see Figure 17.13), allows analysts to determine the location of the beacons with high precision as well as determine the satellite orbits. The system has been placed on a number of satellites intended to carry out geodetic and geophysical research, including Jason-1 and ENVISAT altimetry satellites and the remote sensing satellites SPOT-2, SPOT-4 and SPOT-5. It also flew with SPOT-3 and TOPEX/POSEIDON. As with the other techniques, a service organization called the International DORIS Service has been created to coordinate the observations and data analyses (see Chapter 18) (Tavernier *et al.*, 2005).

17.2.5

Geophysical Modeling

Modern improvements in astronomical, meteorological, and geophysical observations have made it clear that the atmosphere and the oceans do affect the Earth's orientation. This relationship has been explained by analyses of the conservation of angular momentum in the solid Earth, ocean, and atmosphere system. If, for example, a component of atmospheric angular momentum (AAM) increases, the analogous angular momentum component of the solid Earth would decrease to conserve the system angular momentum. This does appear to be the case and the increasing number of accurate environmental observations has led to timely accurate measurements of the angular momentum variations in the atmosphere and ocean that can be compared with the astronomical observations of the Earth's orientation (Koot *et al.*, 2006). Figure 17.14 shows these comparisons for polar motion and the excess length of day.

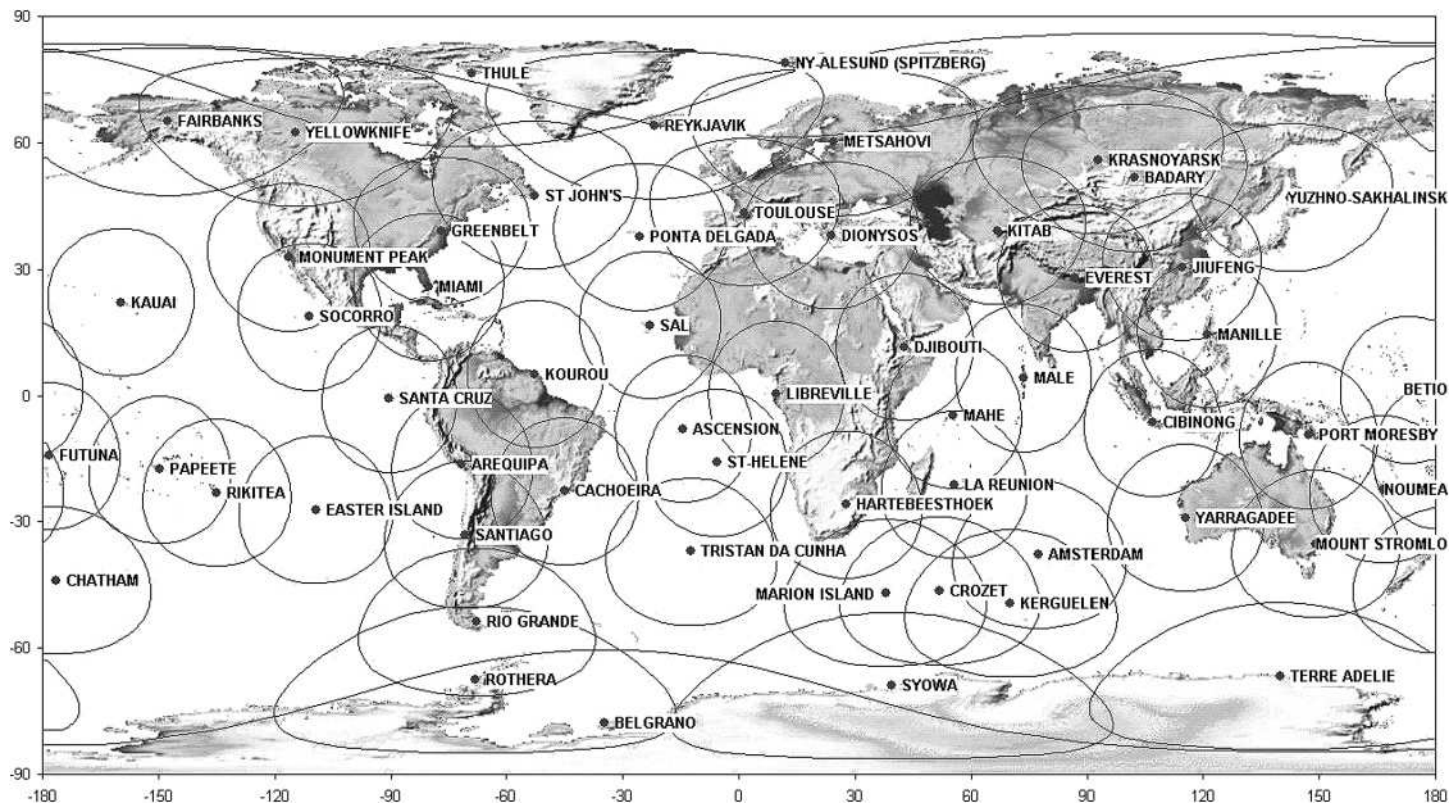


Figure 17.13 DORIS network showing visibilities of satellites with elevations greater than twelve degrees.

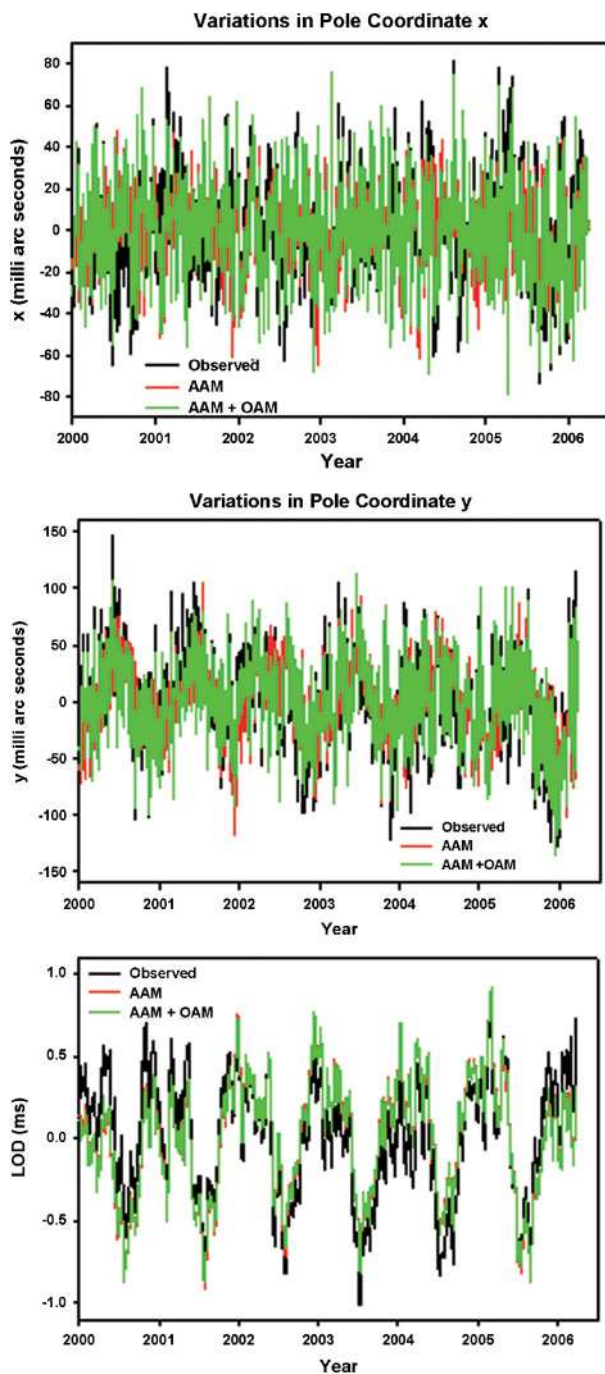


Figure 17.14 Comparison of astronomical observations of polar motion and excess length of day with effects derived from analyses of atmospheric angular momentum (AAM) and oceanic angular momentum (OAM).

It is possible to take advantage of this relationship to improve predictions of the Earth's orientation. Near-term predictions of expected variations in the atmospheric and ocean angular momentum (OAM) can be used to assist in making forecasts of the expected variations in the Earth's orientation. The predictions of the angular momentum of the atmosphere and the ocean are made possible by a network of global environmental observatories and sophisticated mathematical forecast models of these systems. So, while geophysical modeling might not be considered strictly to be a means of observing the Earth's changing orientation, it does make significant contributions to the short-term prediction of polar motion and the Earth's rotation angle.

17.3

Earth Orientation Data

The International Earth Rotation and Reference Systems Service (IERS) (see Chapter 18) routinely provides Earth orientation data produced from the observations available from the sources outlined above. Historical, current, and forecast values of the parameters are made available to users from two product centers, the Earth Orientation Center and the Rapid Service/Prediction Service. The Earth Orientation Center Product Center publishes long-term Earth orientation parameters with the monthly publication of IERS Bulletin B, which lists the most recent values of the Earth's orientation in the IERS Reference System. The IERS Rapid Service/Prediction Center provides Earth orientation parameters on a rapid turnaround basis, primarily for real-time users and others needing the highest quality EOP information sooner than that available in the final series published by the IERS Earth Orientation Center. Table 17.1 lists the uncertainties of the IERS products.

Table 17.1 Uncertainties of Earth orientation parameter observations.

Technique	Sampling time	Precision			
		Polar motion	UT1–UTC	LOD	Celestial pole offsets
VLBI	1 day	0.000 20''	0.000 072 s	0.000 10 s/day	0.000 085''
VLBI	1 hr	–	0.000 250 s	0.000 35 s/day	–
SLR	1 day	0.000 26''	–	0.000 51 s/day	–
GPS	1 day	0.000 12''	–	0.000 35 s/day	–
Combination		0.000 004''	0.000 006 s	0.000 009 s/day	0.000 1''
5-day prediction		0.002''	0.000 6 s	0.000 9 s/day	0.000 1''

References

- Boucher, C., Altamimi, Z., Sillard, P. and Feissel-Vernier, M. (2004) *The ITRF2000*, International Earth Rotation Service Tech. Note 31, Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.
- Carter, W.E. and Robertson, D.S. (1986) Studying the earth by very-long-baseline interferometry. *Sci. Am.*, **255**, 44–52.
- Carter, W.E., Robertson, D.S. and Fallon, F.W. (1989) Polar motion and UT1 time series derived from VLBI observations. *IERS Tech. Notes*, No. 2, 35–9.
- Gurtner, W., Noomen, R. and Pearlman, M.R. (2005) The International Laser Ranging Service: current status and future developments. *Adv. Space Res.*, **36**, 327–32.
- Johnston, K.J. (1979) The application of radio interferometric techniques to the determination of earth rotation, in *Time and the Earth's Rotation* (eds D.D. McCarthy and J.D. Pilkington), Reidel, Dordrecht, pp. 183–90.
- Koot, L., De Viron, O. and Dehant, V. (2006) Atmospheric angular momentum time-series: characterization of their internal noise and creation of a combined series. *J. Geod.*, **79**, 663–74.
- Kouba, J., Beutler, G. and Rothacher, M. (2000) IGS combined and contributed earth rotation parameter solutions, in *Polar Motion: Historical and Scientific Problems*, ASP Conference Series, Vol. 208, also IAU Colloquium #178 (eds S. Dick, D. McCarthy and B. Luzum), ASP, San Francisco, p. 277.
- Ma, C. and Feissel, M. (eds) (1997) *Definition and Realization of the International Celestial Reference System by VLBI Astrometry of Extragalactic Objects*, International Earth Rotation Service Tech. Note 23, Observatoire de Paris, Paris.
- McCarthy, D.D. and Petit, G.P. (eds) (2004) *IERS Conventions (2003)*, (IERS Technical Note 32) Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, p. 127.
- Rothacher, M. (1999) The contribution of GPS measurements to Earth rotation studies, in *Journées 1998–Systèmes de référence spatio-temporels: Conceptual, conventional and practical studies related to Earth rotation*, Observatoire de Paris, pp. 239–47.
- Schlüter, W. and Behrend, D. (2007) The International VLBI Service for Geodesy and Astrometry (IVS): current capabilities and future prospects. *J. Geod.*, **81**, 379–87.
- Schutz, B.E., Tapley, B.D., Eanes, R.J. and Watkins, M.M. (1989) Earth rotation from Lageos laser ranging. *IERS Tech. Notes*, No. 2, 53–7.
- Tavernier, G., Fagard, H., Feissel-Vernier, M., Lemoine, F., Noll, C., Ries, J., Soudarin, L. and Willis, P. (2005) The International DORIS Service. *Adv. Space Res.*, **36**, 333–41.

18

International Activities

18.1

Time and International Activities

Because time has become an international standard, a variety of international organizations have evolved to deal with various aspects of time and timekeeping. These range from the political and commercial agencies concerned with issues related to international standards to scientific organizations dealing with subtle aspects of the technical definitions of time scales and organizations that promote the development of even more accurate devices and means for time dissemination.

18.2

Treaty of the Meter

Although the subject of time was not covered originally in the Treaty of the Meter that was signed on May 20, 1875 (Bureau International des Poids et Mesures, 2006), it has since become part of the mission of the organizational structure put in place by the Treaty. The Treaty of the Meter is also known as the Meter Convention or in French as the Convention du Mètre. Written in the French language, it was signed by 17 countries at the International Metric Convention that was called to organize formally the means to maintain the metric standards. The number of signatories increased to 21 in 1900, 32 in 1950, 44 by 1975, 48 by 1997, and 49 by 2001. As of 2005, there were 51 signatories and 30 states with associate status. It was revised in 1921, and the system of units it established was renamed the *Système international d'unités* (SI) ('International System of Units') in 1960. To carry out the intentions of the treaty, three organizations were created: the Conférence Générale des Poids et Mesures (CGPM), the Comité International des Poids et Mesures (CIPM), and the Bureau International des Poids et Mesures (BIPM). The responsibility for an international standard time was taken on by the CGPM in 1985 (Guinot, 2000).

18.2.1

General Conference on Weights and Measures (CGPM)

Delegates from each of the signatories along with observers from each of the associates comprise the General Conference on Weights and Measures (Conférence Générale des Poids et Mesures, CGPM). The CGPM meets every four years at the BIPM where it receives the official report of the CIPM, discusses possible improvements in the SI units, and endorses new metrological results and international scientific recommendations regarding the fundamental units. It also makes decisions regarding the future direction of the BIPM. The *Système International d'Unités* or the International System of Units, abbreviated 'SI' was established in 1960 by the 11th CGPM, and it is modified by the CGPM as required to reflect the latest advances.

18.2.2

International Committee on Weights and Measures (CIPM)

Eighteen individuals, each from a different member state, comprise the International Committee on Weights and Measures (Comité International des Poids et Mesures, CIPM). Its mission is to promote uniformity in the international measurement units, principally by submitting draft resolutions to the CGPM for its approval. It discusses the work of the BIPM and issues an annual report on the operations of the BIPM to the governments of the member states. Its members discuss and coordinate current metrological activities and prepare other reports including the SI Brochure.

The CIPM has created a number of Consultative Committees (in French: *Comités Consultatifs*) to provide technical information on a wide range of metrological activities. Each committee is composed of technical experts from national metrology institutes, and the chair of each committee usually serves on the CIPM. These committees discuss scientific and technical advances related to metrology and formulate recommendations for the CIPM. They also advise the CIPM on the work of the BIPM. The committees with titles current as of 2008 are:

- Consultative Committee for Acoustics, Ultrasound and Vibration (CCAUV)
- Consultative Committee for Electricity and Magnetism (CCEM)
- Consultative Committee for Length (CCL)
- Consultative Committee for Mass and Related Quantities (CCM)
- Consultative Committee for Photometry and Radiometry (CCPR)
- Consultative Committee for Amount of Substance – Metrology in Chemistry (CCQM)
- Consultative Committee for Ionizing Radiation (CCRI)
- Consultative Committee for Thermometry (CCT)
- Consultative Committee for Time and Frequency (CCTF)
- Consultative Committee for Units (CCU)

The CIPM meets annually at the BIPM to discuss the reports of the Consultative Committees. The CCU assists in the preparation of the SI Brochure. Suggested

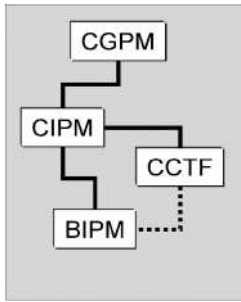


Figure 18.1 Organizational structure within the Treaty of the Meter that deals with time.

modifications of the SI are submitted to the CGPM by the CIPM for formal adoption. On matters relating to interpretation or usage of the SI the CIPM may also adopt its own resolutions and recommendations.

18.2.3

BIPM

The third organization created by the Meter Convention is the Bureau International des Poids et Mesures (International Bureau of Weights and Measures, BIPM). It is located in Sèvres, a suburb of Paris. Its status is that of an intergovernmental organization that is financed by the member states of the Meter Convention. Its operations fall under the supervision of the CIPM (Figure 18.1). The staff of the BIPM, which numbers about 70, carries out its mission to ensure international unification of physical measurements. It is to provide the basis for a single, coherent system of measurements traceable to the SI.

The BIPM is currently organized in five sections:

- Mass
- Time, Frequency, and Gravimetry
- Electricity
- Ionizing radiation
- Chemistry

These carry out a variety of tasks including maintaining the kilogram, coordinating international measurement standards, and, in the case of time, providing the actual SI unit, the second.

18.3

Scientific Unions

International scientific organizations have contributed to accurate timekeeping by promoting investigations of the associated scientific and technical problems. They have made recommendations that have affected modern timekeeping significantly

in the past, and it is expected that they will continue to contribute to future developments. Two scientific unions concerned with precise time, the International Astronomical Union (IAU) and the International Union of Geodesy and Geophysics (IUGG), are members of the International Council for Science (ICSU), a nongovernmental organization of 114 national scientific bodies and 29 international scientific unions. Its activities include planning and coordinating interdisciplinary research to address major issues of relevance in science and society. It serves as an advocate for freedom in pursuing science and promoting access to scientific data and information.

18.3.1

International Astronomical Union (IAU)

The International Astronomical Union (IAU) was founded in 1919 to promote the science of astronomy through international cooperation. It is made up of national and individual members. National members are organizations that represent national professional astronomical communities within their countries, and individual members are professional scientists whose research relates to astronomy. Individual members are elected by the Union's Executive Committee following the recommendation of a National Member.

The IAU is currently organized into 12 divisions. Each division is broken down further into commissions that deal with specific specialized topics. The number of commissions now totals forty. The organization also allows for any number of working groups that can report either to divisions or to commissions. As of 2008 there are 67 national members and over 9600 individual members. The Executive Committee sets and implements the overall policy, and the operations are overseen by a set of elected officers. The center for its business activities is the IAU Secretariat, which is currently hosted by the Institut d'Astrophysique de Paris in France.

In addition to sponsoring a number of symposia each year, the IAU holds a General Assembly every three years. The IAU defines fundamental astronomical and physical constants and astronomical nomenclature. It also promotes educational activities in astronomy and discusses future developments dealing with the science of astronomy. Matters related to the subject of time are discussed in Division 1, which has a number of associated Commissions:

- Commission 4: Ephemerides
- Commission 7: Celestial Mechanics and Dynamical Astronomy
- Commission 8: Astrometry
- Commission 19: Rotation of the Earth
- Commission 31: Time
- Commission 52: Relativity in Fundamental Astronomy

The commissions are composed of technical experts dealing with detailed aspects of the commission's tasks. Commissions 31 and 19 are of particular interest for those dealing with timekeeping. The activities of Commission 31 (Time) include

maintaining cooperation with national and international institutions providing atomic timekeeping information, developing cooperation between observatories and other institutions, providing and archiving astronomical data relevant to atomic timekeeping such as pulsar data, developing methods of analyzing and evaluating astronomical data relevant to fundamental concepts of time, and publicizing astronomical data and results relevant to time. Commission 19 (Rotation of the Earth) supports and coordinates scientific investigations in Earth rotation and related reference frames. Its objectives include encouraging and developing cooperation in observation and theoretical studies of Earth orientation and serving as a link between the astronomical community and those organizations providing the International Terrestrial and Celestial Reference Systems/Frames (ITRS, ITRF, ICRS, and ICRF) and Earth orientation parameters, including the International Association of Geodesy (IAG), International Earth Rotation and Reference System Service (IERS), International VLBI Service for Geodesy and Astrometry (IVS), International GPS Service (IGS), International Laser Ranging Service (ILRS), and International DORIS Service (IDS). It also seeks to develop methods to improve the accuracy and understanding of Earth orientation and related reference systems/frames, ensure agreement and continuity of the reference frames used for Earth orientation with other astronomical reference frames and their densification, and provide the means to compare observational and analysis methods and their results to ensure the accuracy of data and models.

18.3.2

International Union of Geodesy and Geophysics (IUGG)

The International Union of Geodesy and Geophysics (IUGG) is a nongovernmental, scientific organization, established in 1919 to promote international coordination of scientific studies of the Earth and its environment in space. These studies include the shape of the Earth, its gravitational and magnetic fields, the dynamics of the Earth as a whole and of its component parts, the Earth's internal structure, composition and tectonics, the generation of magmas, volcanism and rock formation, the hydrological cycle including snow and ice, all aspects of the oceans, the atmosphere, ionosphere, magnetosphere and solar-terrestrial relations, and analogous problems associated with the Moon and other planets. It is made up of eight semi-autonomous associations, each responsible for specific topics within the Union activities:

- International Association of Cryospheric Sciences (IACS)
- International Association of Geodesy (IAG)
- International Association of Geomagnetism and Aeronomy (IAGA)
- International Association of Hydrological Sciences (IAHS)
- International Association of Meteorology and Atmospheric Sciences (IAMAS)
- International Association for the Physical Sciences of the Ocean (IAPSO)

- International Association of Seismology and Physics of the Earth's Interior (IASPEI)
- International Association of Volcanology and Chemistry of the Earth's Interior (IAVCEI)

These associations can organize individual assemblies in the interim between the IUGG General Assemblies that are held every four years. The IUGG has 65 member countries, most of whom participate in the Union through their national academy or other adhering body.

Most of the activity dealing with timekeeping is carried out through the IAG, which was originally organized as the *Mitteuropäische Gradmessung* (Central European Arc Measurement) in 1862 as the first significant international scientific organization. It became the *Europäische Gradmessung* (European Arc Measurement) in 1867 and in 1886 the *Internationale Erdmessung*. At the first IUGG General Assembly in 1922 it became one of the sections, and it took its present name at the IUGG General Assembly of 1930.

The official mission of IAG is the advancement of geodesy. It is concerned with establishment of reference systems, monitoring the gravity field, Earth rotation, and deformation of the Earth's surface including oceans and ice. It holds symposia and workshops to promote international cooperation and knowledge. It is organized in four commissions:

- Commission 1: Reference Frames
- Commission 2: Gravity Field
- Commission 3: Geodynamics and Earth Rotation
- Commission 4: Positioning and Applications

with an Inter-commission Committee on Theory.

18.3.3

International Telecommunications Union (ITU)

The International Telecommunications Union is a United Nations organization that deals with information and communications technologies. Based in Geneva, Switzerland, the ITU has 191 member states and more than 700 sector members and associates. Sector members are recognized operating agencies, scientific or industrial organizations, and financial or development institutions, and organizations of an international character representing them. The ITU is comprised of three sectors:

- Radiocommunication (ITU-R)
- Standardization (ITU-T)

ITU's standards-making efforts are its best-known—and oldest—activity.

- Development (ITU-D)

Established to help spread equitable, sustainable and affordable access to information and communication technologies (ICT).

The ITU-R manages the international radio-frequency spectrum and satellite orbit resources and is the primary sector of the ITU dealing with issues of time and frequency. The ITU-T deals with setting technical specifications so that elements of communications systems can interoperate seamlessly, and ITU-D creates policies and regulations and provides training programs and financial strategies in developing countries. ITU also organizes TELECOM events that bring together leading elements of the information and communication technologies (ICT) as well as ministers and regulators for exhibitions, and high-level forums.

The ITU dates back to the days following the establishment of telegraph networks. To facilitate international communications countries gradually established regional agreements, and in 1865 the International Telegraph Convention was signed resulting in the formation of the International Telegraph Union. With the development of the telephone and wireless telegraphy it was necessary to establish international agreements regarding radiotelegraphy. The first International Radiotelegraph Conference was held in 1906 in Berlin resulting in the first International Radiotelegraph Convention and a set of regulations. These regulations, which have since been expanded and revised by following radio conferences, are now known as the Radio Regulations.

Within the ITU the International Telephone Consultative Committee (CCIF) was established in 1924, followed by the International Telegraph Consultative Committee (CCIT) in 1925, and the International Radio Consultative Committee (CCIR) in 1927. These organizations coordinated the technical studies, tests and measurements and drew up international standards to ensure international communications. The 1927 International Radiotelegraph Conference was the first to allocate frequency bands to existing radio services, including fixed, maritime and aeronautical mobile, broadcasting, amateur, and experimental.

In 1932 the International Telecommunication Convention was formed by combining the International Telegraph Convention and the International Radiotelegraph Convention. At the same time the name of the Union was changed to International Telecommunication Union to reflect its expanding scope. In 1947 it became a United Nations specialized agency and in 1956, the CCIT and the CCIF were merged to form the International Telephone and Telegraph Consultative Committee (CCITT). In 1992, a plenipotentiary conference, revised the structure of the ITU into the three sectors that integrated the functions carried out by the CCIR and the CCITT.

Some documents of the ITU have the status of international treaties. These are (1) the Constitution and Convention of the International Telecommunication Union originally signed in 1992 and subsequently amended in 1994, 1998 and 2002, and (2) the Administrative Regulations, which include the Radio Regulations (<http://www.itu.int/publ/R-REG-RR/en>) and the International Telecommunication Regulations (<http://www.itu.int/ITU-T/itr/>), which complement the Constitution and the Convention. The last revision of the Radio Regulations was signed 2003, and the International Telecommunication Regulations were signed in 1988. The Radio Regulations incorporate the decisions of the World

Radiocommunication Conferences, including all appendices, resolutions, recommendations and ITU-R recommendations incorporated by reference. World Radiocommunication Conferences (WRC) are held every two to three years to review and, if required, revise the Radio Regulations. The general program of World Radiocommunication Conferences is established four to six years in advance and the final agenda set by the ITU Council two years before the conference, with the concurrence of a majority of Member States.

The Plenipotentiary Conference is held every four years to set the Union's general policies, adopt plans for the future and elect the management team. At this conference ITU member states decide on the future of the organization and sector members can attend as observers. The ITU Council, in the interval between Plenipotentiary Conferences, deals with broad telecommunication policies and prepares a report on the policy and strategic planning of the ITU. It is responsible for ensuring the smooth operation of the Union and facilitates the implementation of the provisions of the ITU Constitution, the ITU Convention, and the Administrative Regulations.

Within the ITU, matters relating to precise time and its dissemination fall within the tasks of the ITU-R, and within the ITU-R they are part of the agenda of Study Group 7. Study Group 7 is part of the structure of the Study Groups which includes:

- Study Group 1 (SG 1)–Spectrum management
- Study Group 3 (SG 3)–Radiowave propagation
- Study Group 4 (SG 4)–Satellite services
- Study Group 5 (SG 5)–Terrestrial services
- Study Group 6 (SG 6)–Broadcasting service
- Study Group 7 (SG 7)–Science services

Within Study Group 7, issues related to precise time fall within the purview of Working Party 7A. Study Group 7 is structured as follows:

- Working Party 7A (WP 7A)–Time signals and frequency standard emissions
- Working Party 7B (WP 7B)–Space Radiocommunication Applications
- Working Party 7C (WP 7C)–Remote Sensing Systems
- Working Party 7D (WP 7D)–Radio Astronomy

Issues related to precise time that might be expected to be included in the Radio Regulations would then be expected to be brought up first with Working Party 7A for discussion. They then would go to Study Group 7, then to ITU-R, before being accepted at a World Radio Conference (see Figure 18.2).

18.4

Service Organizations

Another set of time-related organizations deals with the coordination of observations and the analyses of observations relating to the Earth's rotation. Administra-

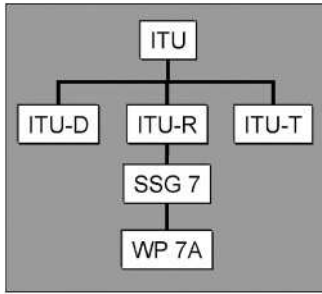


Figure 18.2 Organizational structure within the ITU dealing with time.

tively these organizations fall within the International Council for Science (ICSU), a nongovernmental organization founded in 1931 and composed of 114 national scientific organizations and 29 international scientific unions. Formerly known as the International Council of Scientific Unions, ICSU coordinates interdisciplinary research and provides for the exchange of ideas, the communication of scientific information and the development of scientific standards. It maintains close working relationships with a number of intergovernmental and nongovernmental organizations, especially the United Nations Educational, Scientific and Cultural Organization (UNESCO) and the Third World Academy of Sciences (TWAS). ICSU holds a General Assembly every three years to set general direction, policies and priorities. Its funding is mainly through national members and scientific unions along with grants from UNESCO, the United States, and France.

To address the service aspect related to studies of the Earth, the ICSU formed the Federation of Astronomical and Geophysical Data Analysis Services (FAGS) in 1956. Today it includes twelve permanent services that operate under the authority of one or more of the three scientific unions, International Astronomical Union (IAU), International Union of Geodesy and Geophysics (IUGG), and Union Radio-Scientifique Internationale (URSI). The member services collect observations, information and data related to astronomy, geodesy, geophysics and allied sciences; analyze and distribute data; and publish the results.

Currently the services within FAGS are:

- International Earth Rotation and Reference Systems Service (IERS)
- Bureau Gravimétrique International (BGI)
- International GNSS Service (IGS)
- International Center for Earth Tides (ICET)
- Permanent Service for Mean Sea Level (PSMSL)
- International Service of Geomagnetic Indices (ISGI)
- Quarterly Bulletin of Solar Activity (QBSA)
- International Space Environment Service (ISES)
- World Glacier Monitoring Service (WGMS)
- Center des Données astronomiques de Strasbourg (CDS)
- Solar Influences Data Analysis Center (SIDC)
- International VLBI Service for Geodesy and Astrometry (IVS)

The organization is administered by the FAGS Council composed of representatives from the three scientific unions. Although most of the services touch on areas that affect the Earth's rotation, the services directly related to time within FAGS are the IERS, the IGS, and the IVS. In 2009 the FAGS is expected to merge with the existing World Data Centers to form the World Data System within ICSU.

18.4.1

International Earth Rotation and Reference Systems Service (IERS)

The IERS was established as the International Earth Rotation Service in 1987 by the International Astronomical Union and the International Union of Geodesy and Geophysics, and it began operation on 1 January 1988. In 2003 it was renamed the International Earth Rotation and Reference Systems Service, but retained the acronym IERS. Its objectives are to provide The International Celestial Reference System (ICRS) and its realization, the International Celestial Reference Frame (ICRF), the International Terrestrial Reference System (ITRS) and its realization, the International Terrestrial Reference Frame (ITRF).

The IERS provides Earth orientation parameters required to transform between the ICRF and the ITRF, geophysical data to interpret and model variations in the ICRF, ITRF, or Earth orientation parameters, and the standards, constants and models (i.e., conventions) necessary to use the products. Its products include the International Celestial Reference Frame, the International Terrestrial Reference Frame, monthly Earth orientation data, daily rapid service estimates of near real-time Earth orientation data and their predictions, announcements of the differences between astronomical and civil time for time distribution by radio stations, leap second announcements, products related to global geophysical fluids such as mass and angular momentum distribution, an annual report and technical notes on conventions and other topics, and long-term Earth orientation information.

It carries out its objectives with an organization made up of Technique Centers, Product Centers, Combination Centers, an Analysis Coordinator, and a Central Bureau. The Technique Centers are independent service organizations that have made commitments to the IERS to contribute observational material regarding various aspects of the Earth's rotation. They control the organization of their observations, the analyses and archiving of data, and the development of possible improvements either in the technique or in the analyses of their data. The data are delivered without interruption and with minimal delay. Currently the four Technique Centers are the International GNSS Service (IGS), the International Laser Ranging Service (ILRS), the International VLBI Service (IVS), and the International DORIS Service (IDS). The organization is shown in Figure 18.3, and the details of their individual products are outlined below.

The Product Centers provide the actual products of the IERS. These are various self-supported organizations that have committed to provide these products operationally to the community. The Product Centers include the Earth Orientation Center, the Rapid Service/Prediction Center, the Conventions Center, the International Celestial Reference System Center, the International Terrestrial Refer-

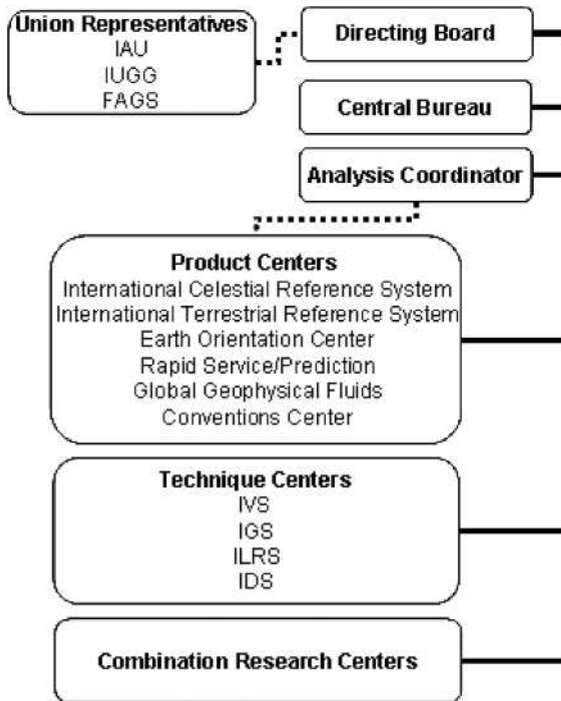


Figure 18.3 Organizational structure of the IERS.

ence System Center, and the Global Geophysical Fluids Center. Within the latter product center, a number of sub-bureaus exist to handle particular aspects of the work. These include the Special Bureau for the Atmosphere, the Special Bureau for the Oceans, the Special Bureau for Tides, the Special Bureau for Hydrology, the Special Bureau for the Mantle, the Special Bureau for the Core, the Special Bureau for Gravity/Geocenter, and the Special Bureau for Loading.

The Earth Orientation Center is situated at the Observatoire de Paris and is responsible for monitoring the long-term aspects of the variations of Earth orientation parameters. It publishes monthly bulletins containing Earth orientation data as well as announcements regarding leap seconds and the difference between UT1 and UTC. The IERS Rapid Service/Prediction Center is provided by the U.S. Naval Observatory and it provides Earth orientation parameters on a rapid turnaround basis, primarily for real-time users and others needing the highest quality EOP information sooner than that available in the final series published by the IERS Earth Orientation Center. It also provides forecasts of future variations in the Earth orientation parameters.

The Conventions Center is a joint operation of the Bureau International des Poids et Mesures and the U.S. Naval Observatory. It is responsible for the maintenance of the IERS conventional models, constants and standards used in the definition and realization of the reference systems. The International Celestial

Reference System (ICRS) Center is responsible for the definition of the ICRS and its realization, the International Celestial Reference Frame (ICRF). This effort is carried out jointly by the Observatoire de Paris and the U.S. Naval Observatory. Similarly the International Terrestrial Reference System (ITRS) Center is responsible for the definition of the ITRS and its realization, the International Terrestrial Reference Frame (ITRF). This work involves network coordination including collocation, local ties, and site quality, and is carried out by the Institut Géographique National (IGN) in France.

The Global Geophysical Fluids Center through its eight Special Bureaus supports research in areas related to the variations in Earth rotation, gravitational field, and geocenter that are caused by mass transport in the geophysical fluids. It is housed at the University of Luxembourg. The Sub-bureaus and their locations are shown in Table 18.1.

IERS Combination Research Centers develop methods to combine data or products provided by different techniques. The ITRS Combination Centers provide ITRF products to the ITRF Product Center after combining inputs from the

Table 18.1 Special Bureaus of the Global Geophysical Fluids Product Center.

Special Bureau	Location	Mission
Atmosphere	Atmospheric and Environmental Research, Inc.	Relevant atmospheric data
Oceans	Jet Propulsion Laboratory	Data relating to nontidal changes in oceanic processes
Tides	NASA Goddard Space Flight Center	Effects of oceanic tides, Earth tides and atmospheric tides
Hydrology	Center for Space Research, University of Texas at Austin	Data sets and numerical models related to the changing distribution of water over the planet
Mantle	Jet Propulsion Laboratory	Information and data on geodynamic effects due to motions in the mantle
Core	Observatoire Royal de Belgique	Data related to the Earth core
Gravity/Geocenter	Jet Propulsion Laboratory	Data on gravity and geocenter
Loading	Nevada Bureau of Mines and Geology and Seismological Laboratory	Products related to surface mass loading

Technique Centers. The Analysis Coordinator is responsible for the long-term internal consistency of the IERS products, and the Central Bureau provides the general administration of the IERS consistent with the Directing Board policies. It is the executive arm of the Directing Board, and it facilitates communications among the components of the organization.

18.4.2

International VLBI Service for Geodesy and Astrometry (IVS)

The International VLBI Service for Geodesy and Astrometry (IVS) is an international collaboration of organizations that provides VLBI data to the IERS and other astrogeodetic users. The objectives of the IVS are to enhance the individual VLBI programs of the member organizations by creating a joint service that coordinates their activities, to promote research and development of the technique, and to serve as an interface with the users of VLBI data. To meet its objectives, the IVS coordinates observing programs, sets standards for observing stations, establishes VLBI data formats, recommends analysis software, and sets up data delivery processes. It carries out its tasks through an organization that includes a Coordinating Center, network stations, operation centers, correlators, data centers, analysis centers, and technology development centers.

The Coordinating Center coordinates both the day-to-day and the long-term activities of the IVS following the directions established by the Directing Board. The network stations are the global VLBI observing sites with geodetic capability that comply with the IVS performance standards for data quality. The operation centers coordinate routine operations of one or more networks. These activities include planning observing programs, establishing operating plans and procedures for the stations in the network, generating the observing schedules, and posting these to an IVS Data Center for distribution and to the Coordinating Center for archiving.

18.4.3

International Laser Ranging Service (ILRS)

The International Laser Ranging Service (ILRS) (Pearlman *et al.*, 2002) is an international collaboration of organizations providing data obtained by laser ranging to artificial Earth satellites or to the Moon. It is one of the services of the IAG. The ILRS collects, archives, and distributes observational data and provides products including polar motion and excess length of day, coordinates and velocities of the ILRS observing sites, time-variable coordinates of the geocenter, static and time-variable models of the Earth's gravity field, satellite ephemerides, lunar ephemerides and librations, and fundamental physical constants. An important user of this information is the IERS, where the data are of particular importance for the maintenance of the ITRF.

The ILRS observational data are obtained by its member tracking stations that range to a constellation of approved satellites and the Moon, using the most

advanced laser tracking equipment. They transmit the ranging data at least daily to one or more operations and/or data centers. These collect and merge the data from other stations in sub-networks, perform data quality checks, and re-format the data as necessary. Global data centers are the primary interfaces to the analysis centers and the outside users and they archive and provide on-line access to tracking data received from the operational/regional data centers. Analysis centers process tracking data from one or more data centers to produce the ILRS products operationally. At a minimum, every analysis center must process the global LAGEOS-1 and LAGEOS-2 data sets and provide Earth orientation parameters on a weekly or sub-weekly basis, as well as other products, such as station coordinates, as required by the IERS. Lunar Analysis Centers process normal point data from the Lunar Laser Ranging (LLR) stations and generate a variety of scientific products including precise lunar ephemerides, librations, and orientation parameters, which provide insights into the composition and internal makeup of the Moon, its interaction with the Earth, tests of general relativity, and solar system ties to the ICRF.

A central bureau coordinates the activities on a daily basis, facilitating communications and information transfer within the organization and with outside users. It maintains the list of satellites approved for tracking support and their priorities, and carries out the directions of an international governing board composed of representatives of the member organizations.

18.4.4

International GNSS (Global Navigational Satellite Service) Service (IGS)

The International GNSS Service (IGS) (Beutler *et al.*, 1999; Dow, Neilan and Gendt, 2005; Kouba *et al.*, 1998) is an international consortium of organizations that provides Global Navigation Satellite Systems (GNSS) tracking data, orbits, and other data products in near real time. The IGS currently offers these products for two GNSS systems, the Global Positioning System (GPS) and GLONASS, and expects to provide similar data in the future for other systems as they become available. It is a service of the International Association of Geodesy and a member of the Federation of Astronomical and Geophysical Data Analysis Services.

IGS products are critical for the improvement and extension of ITRF, monitoring of solid Earth deformations, and monitoring of polar motion. The accuracy of the monitoring station positions is ± 2 mm, and station motion accuracy is ± 3 mm year⁻¹. IGS data are also used to evaluate Earth rotation and variations in sea level, ionospheric monitoring, and measuring precipitable water vapor in the atmosphere. The accuracy of the polar motion data is ± 0.5 mas, and excess length-of-day information is available with an accuracy of ± 0.02 ms day⁻¹.

The IGS accomplishes its objectives with an international network of more than 350 continuously operating dual-frequency GPS monitoring stations, more than a dozen regional and operational data centers, three global data centers, seven analysis centers, and a number of associate or regional analysis

centers. The Central Bureau for the service maintains an Information System that provides access to IGS products. An international Governing Board oversees the services.

The monitoring stations track the satellites with high-accuracy geodetic receivers and send the tracking data to the data centers. They validate and archive the data, making it available to the other elements of the IGS and external users. The analysis centers process the data from the data centers and provide ephemerides, Earth orientation parameters, station coordinates, and clock information. An Analysis Coordinator develops analysis standards and monitors the activities of the analysis centers, providing quality control and performance evaluation. The Analysis Coordinator is also responsible for combining the output of the individual analysis centers into the single set of 'official' IGS products that is sent to the global data centers for distribution. The Central Bureau coordinates the activities and operations of the service consistent with the policies set by an international directing board.

18.4.5

International DORIS Service (IDS)

Like the IVS, ILRS, and IGS, the International DORIS Service (IDS) (Tavernier *et al.*, 2006) provides data from an observational system related to the determination of Earth orientation parameters. DORIS is an acronym for 'Doppler Orbitography and Radio-positioning Integrated by Satellite.' It is a microwave tracking system used to determine the precise location of the satellites on which it is installed. DORIS systems have been installed on the Jason-1 and ENVISAT altimetric satellites and the remote sensing satellites SPOT-2, SPOT-4 and SPOT-5. It also flew with SPOT-3 and TOPEX/POSEIDON. It measures the Doppler frequency shift of a radio signal transmitted from ground stations and received by the satellite. Within the IERS the DORIS data are used primarily for the maintenance and extension of the ITRF.

The IDS is a service of the IAG that collects, archives, and distributes DORIS observations that are then used to determine coordinates and velocities of the ground stations, the geocenter and scale of the Terrestrial Reference Frame, ionospheric information, high-accuracy ephemerides of DORIS satellites, and Earth orientation parameters. Its organization is much like those of the previously described organizations in that it is composed of a network of ground stations, data centers, analysis centers, an analysis coordinator, a central bureau, and a governing board.

The network of ground stations is maintained by the Institut Géographique National (IGN), a French government institution that produces and maintains geographical information for France and its overseas departments and territories. The data centers are also in close contact with the Centre National d'Études Spatiales (CNES), the French government agency responsible for France's space efforts.

References

- Beutler, G., Rothacher, M., Schaer, S., Springer, T.A., Kouba, J. and Neilan, R.E. (1999) The International GPS Service (IGS): an interdisciplinary Service in support of earth sciences. *Adv. Space Res.*, **23**, 631–5.
- Bureau International des Poids et Mesures (2006) *The International System of Units (SI)*, 8th edn, Bureau International des Poids et Mesures.
- Dow, J.M., Neilan, R.E. and Gendt, G. (2005) The International GPS Service (IGS): celebrating the 10th anniversary and looking to the next decade. *Adv. Space Res.*, **36**, 320–6.
- Guinot, B. (2000) History of the Bureau International de l'heure, in *Polar Motion: Historical And Scientific Problems*, Vol. 208 (eds S. Dick, D. McCarthy and B. Luzum), ASP Conference Series. Astronomical Society of the Pacific, San Francisco, pp. 175–84.
- Kouba, J., Mireault, Y., Beutler, G., Springer, T. and Gendt, G. (1998) A discussion of IGS solutions and their impact on geodetic and geophysical applications. *GPS Solut.*, **2**, 3–15.
- Pearlman, M.R., Degnan, J.J. and Bosworth, J.M. (2002) The international laser ranging service. *Adv. Space Res.*, **30**, 135–43.
- Tavernier, G., Fagard, H., Feissel-Vernier, M., Le Bail, K., Lemoine, F., Noll, C., Noomen, R., Ries, J.C., Soudarin, L., Valette, J.J. and Willis, P. (2006) The International DORIS Service: genesis and early achievements, in DORIS Special Issue (ed. P. Willis). *J. Geod.*, **80**, 403–17.

Web Sites of International Organizations

Organization	Web Site
BIPM	www.bipm.org/
FAGS	www.icsuhyphen;fags.org/
IAG	www.iugg.org/associations/iag.html
IAU	www.iau.org/
ICSU	www.icsu.org/
IDS	ids.cls.fr/
IERS	www.iers.org/
IGS	igsb.jpl.nasa.gov/
ILRS	ilrs.gsfc.nasa.gov/
ITU	www.itu.int/
IUGG	www.iugg.org/
IVS	ivscb.gsfc.nasa.gov/

19

Time Applications

19.1

Time Enables the Infrastructure

The infrastructure of modern society depends critically on time and frequency services. We have grown to expect the universal convenience of time and frequency just as we expect the accessibility of electrical power and water services. However, the users of time and frequency services are often not aware of the fact that time and frequency play such important roles in their lives. Requirements for time and frequency with widely varying precision and accuracy exist in the areas of utility services, banking and finance, emergency services, communications, navigation, inventory control, environmental services, transportation management, surveying, agriculture, and recreation, as well as in scientific and technical applications.

19.2

Positioning and Navigation Services

The advent of Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), has not only provided users with an easily accessible source of positioning, navigation, and timing information on a global scale, but it has also imposed new requirements on the providers of time and frequency. Synchronized timing signals are critical to the operation of these systems (see Chapter 17). The fact that a timing error of one nanosecond is roughly equivalent to an error of 30 cm, or one foot, in position means that accurate positioning and navigation requires very precise timing information.

The requirements for timing are, in part, driven by the positioning and navigation applications of the GNSS signals. In addition to the military applications for which GPS was originally intended, these applications now include operation of emergency responders, location of transportation services such as trucks and trains, ship navigation, air traffic control, geodetic land surveys, earthquake monitoring, agricultural and fishing applications, as well as a host of recreational uses.

For example, a developing application for positioning systems is the identification and location of inventory items. Plans are also being considered for the future implementation of an intelligent transportation system (<http://www.its.dot.gov/index.htm>). While these plans are still in the formative stage, it is likely that such a system would impose requirements for navigational errors of less than a few centimeters, implying that timing precision needs better than a tenth of a nanosecond. Land surveying applications and optimization of agricultural operations are further examples of specific applications of improved positioning made possible through precise timing.

Navigation with systems other than GNSS that depend on the timing of electronic signals also requires a level of precision in timing synchronization that is compatible with the system's expected positional accuracy. As an example, the current requirement for LORAN (**LO**ng **R**ange **A**id to **N**avigation) timing is ± 100 ns. This system is a hyperbolic navigation system that provides multiple synchronized low-frequency transmitters so that the user equipment can determine the difference in the time of arrival from pairs of transmitters to determine a position. Future systems such as e-LORAN (enhanced LORAN) can be expected to need timing at the level of ± 5 to a few tens of nanoseconds to meet their navigational objectives.

19.3 Communications

Precise timing is perhaps most critical in communications applications. Common clocks are used, for example, in switching voice and data traffic through the telephone networks. Precise timing is also needed to time stamp information packets and to allow several transmitters to send information simultaneously over a single communication channel. This process, called 'multiplexing,' can be accomplished by means of various techniques. All of them require some form of precise timing or frequency. In this application, it is often the case that, while precision may be critical, accuracy may not be required. This is because a communications network may be operated with local standards for time and frequency, and accuracy is only needed if different networks are required to interoperate with others. At that point, they may require an external standard for time and frequency.

Time division multiplexing (TDM) is a process in which multiple bit streams of information are sent as sub-channels on one communication channel at what appears to be at the same time. In reality the streams are taking turns in time. Fixed time intervals are created so that, for example, a block of data from sub-channel 1 is sent during time interval 1, a block from sub-channel 2 is sent during time interval 2, and so on. The process continues until a frame, which consists of one time interval for each sub-channel, is completed. At that point a new cycle begins with the next block of information from each sub-channel being transmitted. TDM, in the case where several stations are connected to the same physical

medium such as the same frequency channel, is called time division multiple access (TDMA). This process is often used in cellular phone networks, for example. Clearly the time slots must be synchronized at each end of the transmit-receive network to enable the process to work.

Frequency division multiple access (FDMA) divides the radio spectrum into a series of individual frequency bands, allowing each user access to an allocated frequency band without interfering with other users. Multiple-access systems coordinate access between multiple users. Code division multiple access (CDMA) uses spread-spectrum technology and a coding scheme to allow a number of users to share a frequency band. Spectrum management is critical, and that capability depends on accurate frequency standards being available along the communications networks. Each transmitter has an assigned code that is known to the user. The spread-spectrum technique spreads the signal in a particular bandwidth over a frequency domain. To reconstruct the signal at the receiver, the transmitter and receiver must use the same coding scheme to spread the signal over the frequency domain.

The American National Standards Institute (ANSI) describes telecommunications timing requirements by means of four 'stratum' levels. According to the ANSI standard 'Synchronization Interface Standards for Digital Networks' (ANSI, 1987, T1.101-1987), a Stratum 1 timing source is an autonomous source of timing requiring no other input than possibly a yearly calibration. It must perform with a maximum drift rate, defined in terms of fractional frequency $\Delta f/f$, better than $\pm 1 \times 10^{-11}$ over a year. A properly calibrated Stratum 1 source, then, is capable of providing bit-stream timing that will not slip relative to a perfect standard more than once every 4 to 5 months. When a frame slip does occur in network voice equipment, frame synchronization is generally quickly re-acquired, resulting in an audible pop or click, but data circuits lose a number of bits depending on the data rate being transmitted and on whether or not an error correction protocol is being used. A typical stratum 1 timing source is often an atomic standard (cesium beam or hydrogen maser) or an oven-controlled quartz crystal oscillator, even though atomic standards are capable of better accuracy than these specifications. A properly calibrated clock system controlled by means of a GPS timing signal may also be considered for use as a Stratum 1 timing source. In telecommunications networks, Stratum 1 sources are considered to be the primary sources of time and frequency for networks.

Stratum 2 sources must be capable of being adjusted to match the frequency of the Stratum 1 source within a range of $\pm 1.6 \times 10^{-8}$ in one year and drift no more than $\pm 1 \times 10^{-10}$ in a day. This will provide a frame slip rate of about 1 slip in 7 days when the Stratum 2 source is not being updated by a Stratum 1 source. Typical Stratum 2 sources are rubidium standards and oven-controlled quartz oscillators. Table 19.1 compares all of the stratum level definitions.

Stratum 3E was developed after Stratum 3 was standardized in order to meet SONET (synchronous optical networking) equipment requirements. SONET is a protocol that is used to transfer multiple digital bit streams over an optical fiber using lasers or light-emitting diodes (LEDs). Stratum 4 sources are designed to

Table 19.1 Standards for stratum levels.

Stratum	Adjustment accuracy	Maximum drift rate	Time between frame slips
1	1×10^{-11}	—	72 days
2	1.6×10^{-8}	$1 \times 10^{-10}/\text{day}$	7 days
3E	4.6×10^{-6}	$1 \times 10^{-8}/\text{day}$	3.5 hours
3	4.6×10^{-6}	$3.7 \times 10^{-7}/\text{day}$	6 minutes
4	32×10^{-6}	$32 \times 10^{-6}/\text{day}$	—

track Stratum 2 or 3 sources and have no holdover capability. They are not recommended as a timing source for any other system outside the network.

In Network Time Protocol (NTP) applications, the term ‘stratum level’ is used differently. For these purposes stratum level refers to the distance of a network server from a reference clock. Thus, a Stratum 0 source is assumed to be an accurate source of time and frequency with minimal delays which cannot be used on the network. Such sources can be connected to computers, which act as Stratum 1 servers. A Stratum 1 time server must be directly linked to a reliable source of UTC, not over a network path, and acts as a primary network time standard. A Stratum 2 server gets its time over a network link, using NTP, from a Stratum 1 server, and a Stratum 3 server gets its time over a network link, using NTP, from a Stratum 2 server, and so on.

19.4

Power Grid

Precise timing becomes a concern for the distribution of electrical power in more than one specific application. The efficiency of transmitting power throughout grids depends on precise matching of the phases of the alternating electrical current. This requires an accurate frequency reference across the network of power grids. Typically this phase information is obtained by individual power companies by maintaining local sources of time that are referred to a standard timing signal. GPS supplies a cheap, easy-to-use timing signal that provides time with an accuracy orders of magnitude better than that necessary for phase matching. For example, a \pm one-degree phase difference in a 60 Hz oscillation corresponds to a $\pm 46 \mu\text{s}$ time difference, so that maintaining phase matching at the level of $\pm 0.1^\circ$ would require timing at the level of $\pm 5 \mu\text{s}$. In general, adequate capability can be maintained using rubidium standards and quartz oscillators that are steered to a GNSS timing signal. Temporary loss of that signal could be mitigated by the holdover capacity of the time and frequency standards and their backups.

A second application of precise timing in electrical power applications is in locating faults or breaks in the grid. A $\pm 1 \mu\text{s}$ timing capability is adequate to meet these applications. Other applications of timing information deal with billing and

secondary aspects of power distribution. These requirements lie in the range of ± 0.1 s to ± 1 ms.

19.5

Banking and Finance

Timing information with relatively low accuracy is used by banking and finance institutions to time stamp financial transactions. Generally, an accuracy of only three seconds is needed for time-stamping transactions. For this purpose a timing system with accuracy of ± 0.1 s is more than adequate and easily achievable by a number of means. Typically, timing networks are used that involve rubidium or quartz clocks that are steered using NTP or GPS. These clocks would be capable of relatively long periods without being steered in the event of the loss of an outside source of timing information.

19.6

Emergency Services

Timing requirements for time-stamping standard emergency service operations are about the same as those for the banking and finance sector. Emergency service centers generally rely on radio services and NTP as sources for accurate time, with backup systems of rubidium or quartz clocks. Enhanced emergency services are being developed in which callers can be located through the use of GNSS services. This capability promises faster responses and improved geo-location of required services. It is essentially an application of existing GNSS positioning, and, as such, the timing requirements are equivalent to those discussed in Section 19.2.

19.7

Water Flow

Precise timing information is even used for time stamping by water utilities to control flow in and out of local systems through remotely controlled valves and to control the chemical treatment of water and sewage. As with banking and finance applications, the accuracy required is relatively low and easily achievable using widely available sources of time and low-precision clocks.

19.8

Summary

Table 19.2 provides an outline of the applications of precise time and frequency along with representative accuracy and precision needs. Some need only precision,

Table 19.2 Representative estimates of time and frequency needs.

Community	Application	Purpose	Accuracy		Precision	
			Time	Frequency	Time	Frequency
Positioning and Navigation	Aviation	Fuel management, traffic spacing, course navigation	±3 ns	$\pm 3.5 \times 10^{-14}$		
	Space	Artificial satellite location, surveillance	±25 ns	$\pm 3.0 \times 10^{-13}$		
	Maritime	Fuel management, routing, cargo location	±25 ns	$\pm 3.0 \times 10^{-13}$		
	Transportation	Fuel management, real-time routing, cargo location, intelligent transportaion system	±1 ns	$\pm 1.0 \times 10^{-14}$		
	Agriculture	Field management, fertilizer optimization, livestock tracking	±10 ns	$\pm 1.0 \times 10^{-13}$		
	Railroad	Asset location, real-time routing	±10 ns	$\pm 1.0 \times 10^{-13}$		
	Automotive	Intelligent highway system	±1 ns	$\pm 1.0 \times 10^{-14}$		
	Recreation	Small boat navigation, hiking, bicycle touring	±25 ns	$\pm 3.0 \times 10^{-13}$		
Communications	Voice	Cellular phones			±1.0 ns	$\pm 1.0 \times 10^{-14}$
	Data	Data transmission, secure communications			±1.0 ns	$\pm 1 \times 10^{-11}$

Survey and Mapping	Geographic Information Systems	Asset location	$\pm 10 \text{ ns}$	$\pm 1.0 \times 10^{-13}$		
	Datum Management	Land management	$\pm 25 \text{ ns}$	$\pm 3.0 \times 10^{-13}$		
Energy	Power Grid	Phase matching	$\pm 50 \mu\text{s}$	$\pm 6.0 \times 10^{-10}$		
	Oil & gas location	Exploration	$\pm 25 \text{ ns}$	$\pm 3.0 \times 10^{-13}$		
	Power Line management	Fault location			$\pm 10 \text{ ns}$	$\pm 1.0 \times 10^{-13}$
Banking and Finance	Asset management	Time stamping	$\pm 0.1 \text{ s}$	$\pm 1.0 \times 10^{-6}$		
Emergency Services	Fire, police, medical services	Search and rescue, vehicle location	$\pm 10 \text{ ns}$	$\pm 1.0 \times 10^{-13}$		
Water	System operation	Flow management			$\pm 0.1 \text{ s}$	$\pm 1.0 \times 10^{-6}$
Environmental	Resource management	Hazardous waste containment	$\pm 25 \text{ ns}$	$\pm 3.0 \times 10^{-13}$		
Scientific	Geodesy	Plate tectonics, ocean level	$\pm 1.0 \text{ ns}$	$\pm 1.0 \times 10^{-14}$		
	Astronomy	Pulsar investigations	$\pm 1.0 \text{ ns}$	$\pm 1.0 \times 10^{-14}$		
	Physics	Measurement precision	$\pm 0.01 \text{ ns}$	$\pm 1.0 \times 10^{-16}$		

but others require accuracy or the ability to be traceable to a conventional standard. The table shows that many users of precise time and frequency information do not require the ultimate in either quantity and that there is a very large range in the needs of users. Consequently the relative cost of the apparatus is likely to drive the choices of equipment used in the various applications.

Reference

ANSI (1987) T1.101-1987, *Synchronization Interface Standards for Digital Networks*, American National Standards Institute, New York.

20

Future of Timekeeping

20.1

Future Needs for Time

The demand for improvements in the precision and accuracy of our measurement of time and frequency will drive the developments in the field for years to come. The applications on the horizon that would make use of this improved capability are as varied as the current applications.

Efficient means to improve the transportation methods of the future are just one of those applications. More accurate positioning of aircraft, made possible through improved synchronization of future timing standards, will enable more efficient use of air space. This might mean that not only could aircraft be spaced more efficiently, but that air traffic could make use of more direct flight paths rather than being constrained as they are currently. Plans are under way for managing ground transportation in a similar way. This could mean that at some future time drivers could take advantage of a managed network of roads controlled through precise navigation of vehicles. The efficiency of such a system will depend on precise positioning, again made possible by synchronized timing signals spaced along the highways of the future.

Currently the inability to navigate indoors, or in areas where Global Navigation Satellite System (GNSS) signals are too weak to be effective, limits the effectiveness of responses to emergency situations. It is likely that building planners of the future will consider indoor navigation capabilities, enabled by precise time signals within buildings. Plans to take advantage of existing GNSS signals to aid in locating emergency situations are already being implemented. These types of services are likely to be improved in the future as timing capabilities improve.

In the future the management of resources will be able to take advantage of accurate location of inventory items. Identification of items and their whereabouts will be an important tool available to managers. The usefulness of such identification and location systems will depend directly on the resolution precision of those systems. This capability will again depend on precise positioning made possible by improved timing capabilities.

The communications facilities of the future may be the most direct beneficiaries of improved time and frequency. Currently systems are operating at the edge of our capability to provide sufficient bandwidth. More precise spectrum allocation made possible by improved frequency standards will enable a greater throughput of data in the future.

The scientific community will also benefit significantly from enhanced time and frequency capability. Measurements depending on accurate time or precise frequency resolution will be instrumental in future developments. This should benefit the fields of geodesy, physics, astronomy, Earth dynamics, and optical measurements, to mention just a few fields which are affected by the accuracies of time and frequency.

20.2

Modeling the Earth's Rotation

Future improvements in Earth rotation modeling are likely to depend on both improved observations and improvements in understanding the physics of the Earth. In the area of astronomical observations of the Earth's orientation we can expect observations to be made with higher time resolution, so that observations of the Earth's rotation vector in space and in the terrestrial frame could be available routinely with sub-daily frequency. We can also expect that observations of the phenomena known to affect the rotation vector will also be made available with higher resolution in both time and space. For example, observations of winds and ocean currents are likely to be available from observing platforms in space and from a dense grid of terrestrial observing locations. This information will conceivably lead to estimates of atmospheric and ocean angular momentum with hourly frequency.

Improved modeling of lower frequency variations in the Earth's rotation will depend on improved understanding of the physical processes of the Earth's interior. Advances in the measurement of the gravitational and magnetic fields could be helpful in this endeavor. The understanding of the variations in these fields, in both time and space, is likely to improve knowledge of the response of the Earth's rotation vector to the physical phenomena occurring within the Earth. This knowledge may lead to understanding of the drivers of polar motion and the forces changing the Earth rotation.

Using these observations, together with improved capacity to analyze the data quickly, we might expect true real-time estimates of the Earth's rotation angle in space as well as nearly real-time values of the variations in the Earth's rotational speed. We may even get to the point that we could predict more reliably the Earth's rotation for months in advance, but that capability will continue to be a challenge for the future.

20.3

Clocks of the Future

Future clocks will be available to cover a wide range of users' needs. Chip-scale atomic clocks are becoming readily available now, and we can imagine a future when individuals interested in extremely precise time could be wearing watches that offer time based on atomic energy level transitions. Assuming that such devices can be made sufficiently robust, these clocks are likely to be adapted to a number of applications depending on user requirements.

Optical clock technology is likely to be extended to provide the future laboratory standards. We can expect developments of spectroscopic techniques that will make available extremely high- Q transitions for the laboratory standards of the future. These will likely be augmented by fountains using alkali metal atoms, and will gradually replace the cesium beam oscillators of today.

The necessity for development of improved devices will come from the need for improvements in positioning and navigation, as well as the growing requirements for the transfer of information. Expectations for improved communications capabilities continue to grow. These expected improvements include higher data rates and more efficient use of the radio frequency spectrum in order to provide greater information throughput. All of these applications are likely to take advantage of improvements in frequency standards.

The use of pulsars to provide long-term frequency stability to current time scales has been investigated in the past, but so far has not been found to be useful operationally. We might expect that the investigation of the physics of pulsars will lead to a better understanding of their rotational speeds. That might lead to a need for an ensemble of accurately measured millisecond pulsars to form a statistical base for a time scale and to identify changes in the frequencies of individual pulsars. This may then make the use of pulsars more attractive for their contribution as 'clocks' to the stability of the time scales of the future.

20.4

Future Time Scales

Time scales can be expected to keep pace with the development of the future time and frequency standards. This will not happen without development of statistical models to evaluate the performance of contributing clocks and the development of robust algorithms to combine the best features of the contributors to the time scale. The accuracy of clock comparisons will need to be developed to make this process viable, and for this purpose, improved calibration techniques will need to be developed.

We can expect that the future of leap seconds as a means to reconcile the use of atomic time with observations of the Earth's rotation will be resolved in the near future. The need for a time scale free of unpredictable one-second adjustments in

epoch is clear. Likewise, there is the continuing need for a means to provide the specific information regarding the Earth's rotation, as given now by UT1. The future will decide if a leap-second-less time scale will be in the realm of the everyday user, or if the users of time and frequency will prefer two time scales: one for 'everyday' use related to the Earth's rotation and another for use in precise time applications without leap seconds. The improved accuracies may provide a new basis for the commonly available time scales. In any event we can expect that the future time scales will continue to be based on atomic energy level transitions.

The SI second of today's time scales is defined by the hyperfine transition of the cesium atom. However, this may become problematic for the future definition of the second. As new time and frequency standards are developed that take advantage of higher- Q spectral lines, the current definition of the second may no longer be adequate. It just might not provide the accuracy needed for future developments. Consequently, we might expect that the second could be re-defined in some way so as to be compatible with the past and still have the precision needed for future applications. Many scientific and technological standards today (e.g., for the meter) are based on the current definition of the second, so it is difficult to imagine a revised definition of the second that would not be compatible with the past.

In the future we might expect operationally standard time scales very similar to the atomic time scales of today, that is, UTC and/or TAI. We can also expect that the time scales implicit in the global navigation satellite systems of the future will play a larger role in everyday usage. We can only hope that these scales will not differ significantly from the international standards and produce 'nonstandard' time scales.

A growing concern for future time scales will be the need for an easily realizable time scale for use in space. We currently define such time scales, for example, TCB, TCG, and TDB, for use in dynamical applications in the solar system. With the expected improvements in accuracies, the family of dynamical time scales will need to have better definitions based on higher orders of relativity and the post-Newtonian parameters. While these scales are necessary for the development of ephemerides and reduction of planetary and lunar observations, they are not used for actual operations in space missions. For this purpose we will need relativistically correct time scales easily comparable to clocks on the Earth for future planetary expeditions. The timing accuracy required for navigation and communications for missions at Mars indicates the need for an accurate relation between UTC and a Martian proper time.

20.5

Future Time Distribution

With all of the developments likely to occur as outlined above, questions remain about how we might expect to get time to the user of the future. There is no reason to expect that the future satellite navigation systems will not continue to provide

precise time to the vast majority of precise time users. We might also expect that time provided by those systems will be limited only by the ability to calibrate the user's receiver. Those needing time and frequency with qualities similar to those provided by laboratory standards are likely to make use of optical fibers. This technology, which is available today, is likely to be extended in the future, and we might imagine a future when the use of optical fibers for time and frequency applications is ubiquitous.

An intriguing possibility that has been discussed is that of quantum entanglement. This is a theoretical phenomenon in which objects that are separated in space have quantum states that are linked in such a manner that when one changes state the other must also change. Should this be true, it would have obvious implications for instantaneous time synchronization. However, many have argued that information cannot be transferred faster than the speed of light, so that this technique would not be a viable means to transfer time.

Another possibility more likely for development in the future is that of a distributed clock. In this system, the clock that is to be used as the conventional standard will actually be made up of a large number of interrelated clocks both on the Earth's surface and in space. The relationship between these clocks might be known through optical fiber connections or laser ranging connections. If the potential bias and rate of any clock in this distributed system is known nearly instantaneously, then that clock provides a nearly instantaneous traceability to the conventional standard. Such a system will depend critically on our ability to calibrate the time comparison network of the future.

Glossary

ΔT : the difference between **Terrestrial Time** and **Universal Time**, specifically the difference between Terrestrial Time (TT) and UT1: $\Delta T = TT - UT1$.

$\Delta UT1$ (or ΔUT): the value of the difference between UT1 (UT) and UTC.

aberration: the apparent angular displacement of the observed position of a celestial object from its **geometric position**, caused by the finite velocity of light in combination with the motions of the observer and of the observed object. Annual aberration is due to the motion of the Earth around the Sun, while diurnal aberration is due to the Earth's rotation.

accuracy: closeness of an estimated (e.g., measured or computed) value to a standard or accepted value of a particular quantity.

Allan variance: the square root of the sum of the squares of the differences between consecutive readings divided by twice the number of differences.

almanac, astronomical: an annual publication containing information on the locations of celestial bodies, together with the times and circumstances of various astronomical events such as sunset and sunrise, of particular use for navigation.

altitude: the angular distance of a celestial body above or below the horizon, measured along the great circle passing through the body and the zenith. Altitude is 90° minus zenith distance.

analemma: a curve showing the angular offset of a celestial body (usually the Sun) from its mean position on the celestial sphere as viewed from another celestial body (usually the Earth). In the case of the Sun as seen from the Earth this is a curve resembling a figure of eight that is commonly printed on globes.

angular momentum: measure of the extent to which an object will continue to rotate about a reference point unless acted upon by an external torque. Angular momentum is related to the mass, velocity and distance of an object from the reference point.

anomaly: angular measurement of a body in its orbit from its pericenter.

aphelion: the point in a planetary orbit that is at the greatest distance from the Sun.

apogee: the point at which a body in orbit around the Earth reaches its farthest distance from the Earth. Apogee is sometimes used with reference to the apparent orbit of the Sun around the Earth.

apparent place: the position at which the object would actually be seen from the center of the Earth, displaced by planetary aberration (except the diurnal part) and referred to the true equator and equinox.

apparent right ascension and declination: angular coordinates in the true equator and equinox of date reference system at a specified date. They are geocentric positions differing from the ICRS positions by annual parallax, gravitational light deflection due to the solar system bodies except the Earth, annual aberration, and the time-dependent rotation describing the transformation from the GCRS to the Celestial Intermediate Reference System (they are similar to intermediate positions in the CIO based system but the apparent right ascension origin is at the equinox). Note that apparent declination is identical to intermediate declination.

apparent solar time: the measure of time based on the diurnal motion of the true Sun. The rate of diurnal motion undergoes seasonal variation because of the obliquity of the ecliptic and the eccentricity of the Earth's orbit. Additional small variations result from irregularities in the rotation of the Earth on its axis.

aspect: the apparent position of any of the planets or the Moon relative to the Sun, as seen from Earth.

asterism: a pattern of stars seen in the Earth's sky which is not an official constellation.

astrometric ephemeris: an ephemeris of a solar system body in which the tabulated positions are essentially comparable to catalog mean places of stars at a standard epoch. An astrometric position is obtained by adding to the geometric position, computed from gravitational theory, the correction for light-time.

astronomical coordinates: the longitude and latitude of a point on Earth relative to the geoid. These coordinates are influenced by local gravity anomalies.

astronomical unit (a.u.): distance from the center of the Sun at which a particle of negligible mass, in an unperturbed circular orbit, would have an orbital period of 365.256 898 3 days. This is slightly less than the semimajor axis of the Earth's orbit. The precise value is currently accepted as $149\,597\,870\,691 \pm 30$ meters.

atomic second: *see* second, *Système International*.

axis of inertia: the axis of a principal moment of inertia. In the case of Earth, if it is considered to be symmetrical under rotation about a given axis, the symmetry axis is a principal axis. Also referred to as the axis of figure.

axis of rotation: the instantaneous axis about which the Earth's rotation is taking place. It moves slowly around the axis of figure and this motion is referred to as polar motion.

azimuth: the angular distance measured clockwise along the horizon from a specified reference point (usually north) to the intersection with the great circle drawn from the zenith through a body on the celestial sphere.

barycenter: the center of mass of a system of bodies; for example, the center of mass of the solar system or the Earth-Moon system.

Barycentric Celestial Reference System (BCRS): a system of barycentric space-time coordinates for the solar system within the framework of General Relativity with metric tensor specified by the IAU. Formally, the metric tensor of the BCRS does not fix the coordinates completely, leaving the final orientation of the spatial axes undefined. However, for all practical applications the BCRS is assumed to be oriented according to the ICRS axes.

Barycentric Coordinate Time (TCB): the coordinate time of the BCRS; it is related to Geocentric Coordinate Time (TCG) by relativistic transformations that include secular terms.

Barycentric Dynamical Time (TDB): a time scale originally intended to serve as an independent time argument of barycentric ephemerides and equations of motion. In the IAU 1976 resolutions, the difference between TDB and TDT was stipulated to consist of only periodic terms, a condition that cannot be satisfied rigorously. The IAU 1991 resolutions introducing barycentric coordinate time (TCB) noted that TDB is a linear function of TCB, but without explicitly fixing the rate ratio and zero point, leading to multiple realizations of TDB. In 2006 TDB was re-defined through a linear transformation of TCB.

Barycentric Ephemeris Time (T_{eph}): the time scale of the JPL DE ephemeris that has been scaled to Terrestrial Time. It is equivalent to the redefined TDB.

Cesium fountain: an apparatus that realizes the SI definition of the second by vertically launching Cesium atoms through a microwave cavity and allowing gravity to bring the atoms back down through the cavity.

calendar: a system of reckoning time in which days are enumerated according to their position in cyclic patterns.

catalog equinox: the intersection of the hour circle of zero right ascension of a star catalog with the celestial equator.

Celestial Ephemeris Origin (CEO): the original name for the Celestial Intermediate Origin (CIO) given in the IAU 2000 resolutions.

Celestial Ephemeris Pole (CEP): the reference pole for nutation and polar motion used from 1984 to 2003 with the IAU 1980 Theory of Nutation; the axis of figure for the mean surface of a model Earth in which the free motion has zero amplitude. This pole was originally defined as having no nearly-diurnal nutation with

respect to a space-fixed or Earth-fixed coordinate system and being realized by the IAU 1980 nutation. It is now replaced by the CIP.

celestial equator: the plane perpendicular to the Celestial Ephemeris Pole. Colloquially, the projection onto the celestial sphere of the Earth's equator.

Celestial Intermediate Origin (CIO): origin for right ascension on the intermediate equator in the Celestial Intermediate Reference System. It is the nonrotating origin in the GCRS that is recommended by the IAU. The CIO was originally set close to the GCRS meridian and throughout 1900–2100 stays within 0.1 arcseconds of this alignment.

Celestial Intermediate Pole (CIP): geocentric equatorial pole defined as being the intermediate pole, in the transformation from the GCRS to the ITRS, separating nutation from polar motion. It replaced the CEP on 1 January 2003. Its GCRS position results from (i) the part of precession-nutation with periods greater than 2 days, (ii) the retrograde diurnal part of polar motion (including the free core nutation, FCN), and (iii) the frame bias. Its ITRS position results from (i) the part of polar motion which is outside the retrograde diurnal band in the ITRS and (ii) the motion in the ITRS corresponding to nutations with periods less than 2 days. The motion of the CIP is realized by the IAU precession-nutation plus time-dependent corrections provided by the IERS.

Celestial Intermediate Reference System (CIRS): geocentric reference system related to the GCRS by a time-dependent rotation taking into account precession-nutation. It is defined by the intermediate equator (of the CIP) and CIO on a specific date. It is similar to the system based on the true equator and equinox of date, but the equatorial origin is at the CIO.

celestial pole: either of the two points projected onto the celestial sphere by the extension of the Earth's axis of rotation to infinity.

celestial pole offsets: time-dependent corrections to the precession-nutation model, determined by observations. The IERS provides the celestial pole offsets in the form of the differences, dX and dY , of the CIP coordinates in the GCRS with respect to the IAU 2000A precession-nutation model (i.e., the CIP is realized by the IAU 2000A precession-nutation plus these celestial pole offsets). In parallel the IERS also provides the offsets, $d\Psi$ and $d\epsilon$, in longitude and obliquity with respect to the IAU 1976/1980 precession/nutation model.

celestial pole offsets at J2000.0: offset of the direction of the mean pole at J2000.0, provided by the current model, with respect to the GCRS. These offsets are part of what is often called frame bias.

celestial sphere: an imaginary sphere of arbitrary radius upon which celestial bodies may be considered to be located. As circumstances require, the celestial sphere may be centered at the observer, at the Earth's center, or at any other location.

center of figure: that point so situated relative to the apparent two-dimensional figure of a body that any line drawn through it divides the figure into two parts having equal apparent areas. If the body is oddly shaped, the center of figure may lie outside the figure itself.

Chandler wobble: the approximately 435-day periodic motion of the CIP in the ITRF corresponding to the free motion of the nonrigid Earth.

chronometer: high-precision, portable timekeeping device.

clepsydra: a device to measure time based on the uniform flow of water. From the Greek, literally it means a water thief.

clock: any device for indicating the time.

conjunction: the phenomenon in which two bodies have the same apparent celestial longitude or right ascension as viewed from a third body. Conjunctions are usually tabulated as geocentric phenomena. For Mercury and Venus, geocentric inferior conjunction occurs when the planet is between the Earth and Sun, and superior conjunction occurs when the Sun is between the planet and Earth.

constellation: a grouping of stars, usually with pictorial or mythical associations, that serves to identify an area of the celestial sphere. Also, one of the precisely defined areas of the celestial sphere, associated with a grouping of stars that the IAU has designated as a constellation.

Conventional International Origin: the international origin of polar motion adopted for use by the former International Latitude Service (ILS). It was defined in 1967 by an adopted set of astronomical latitudes of the 5 stations of the ILS. It approximately coincided with the mean pole of 1903.0 as determined by the ILS. To avoid ambiguity, this origin should be designated by its full name. This designation should be avoided for the current origin (the ITRF pole) of the polar motion, which no longer coincides with the conventional international origin.

coordinate time: the time coordinate, which, together with three spatial coordinates specify an event in a four-dimensional space-time reference system. An unambiguous way of dating and the time based on the theory of motion in a specific reference system.

Coordinated Universal Time (UTC): the time scale differing from TAI by an integral number of seconds; it is maintained within ± 0.90 second of UT1 by the introduction of one second steps (leap seconds).

Coriolis effect: an apparent deflection of moving objects when they are viewed from a rotating frame. The Coriolis effect is caused by a fictitious Coriolis force, which appears in the equation of motion of an object in a rotating frame of reference.

Coulomb interaction: the static interaction between two charges, Q_1 and Q_2 , separated by a distance r producing a force proportional to $Q_1 Q_2/r^2$.

crustal motion: the motion of large surface areas of the Earth with respect to other surface areas, also referred to as plate tectonics and continental drift.

culmination: passage of a celestial object across the observer's meridian; also called 'meridian passage.' More precisely, culmination is the passage through the point of greatest altitude in the diurnal path. Upper culmination (also called 'culmination above pole' for circumpolar stars and the Moon) or transit is the crossing closer to the observer's zenith. Lower culmination (also called 'culmination below pole' for circumpolar stars and the Moon) is the crossing farther from the zenith.

day: an interval of 86 400 SI seconds, unless otherwise indicated.

decadal variations: quasi-periodic components of polar motion or observed variations of the Earth's rotation with periods between two and forty years apparently due to interactions of the Earth's mantle and liquid core. Also decadal polar motion, decadal irregularities, decadal fluctuations.

decans: groupings of stars used to reckon time in ancient Egypt. They were made up of 36 asterisms.

declination: angular distance on the celestial sphere north or south of the celestial equator. It is measured along the hour circle passing through the celestial object. Declination is usually given in combination with right ascension or hour angle.

deferent: in representing the motion of planets with epicycles, the circle centered around a point halfway between a point called the equant and the Earth on which an epicycle moves with uniform motion, not with respect to the center, but with respect to the equant.

deflection of light: the angle by which the apparent path of a photon is altered from a straight line by the gravitational field of a massive object situated along the path of the photon. In the case of the Sun the path is deflected radially away from the Sun by up to 1.75'' at the Sun's limb. Correction for this effect, which is independent of wavelength, is included in the reduction from mean place to apparent place.

deflection of the vertical: the angle between the astronomical and the geodetic vertical

delta T: *see* ΔT .

delta UT1: *see* $\Delta UT1$.

direct motion: for planetary movement in the solar system, motion that is counterclockwise in the orbit as seen from the north pole of the ecliptic; for an object observed on the celestial sphere, motion that is from west to east, resulting from the motion of the object relative to the Earth.

diurnal motion: the apparent daily motion, caused by the Earth's rotation, of celestial bodies across the sky from east to west.

Doppler cooling: a mechanism based on the Doppler Effect used to trap and cool atoms, and sometimes used synonymously with laser cooling.

Doppler Effect: the change in frequency of a received signal due to the relative motion of an observer with respect to the source of the signal.

DUT1: predicted value $UT1 - UTC$ to an accuracy of 0.1 s.

dynamical equinox: the ascending node of the Earth's mean orbit on the Earth's true equator; that is, the intersection of the ecliptic with the celestial equator at which the Sun's declination changes from south to north.

dynamical time: the independent variable of the equations of motion of solar system bodies.

Earth's inner core: a solid sphere of nickel and iron at the center of the Earth about 1220 km in radius.

Earth's mantle: a highly viscous layer directly under the crust and above the outer core, about 2900 km thick.

Earth orientation: information specifying the relationship of terrestrial and celestial reference frames.

Earth's outer core: liquid layer, approximately 2300 km thick, composed of iron and nickel above the solid inner core.

Earth Rotation Angle (ERA): angle measured along the intermediate equator of the Celestial Intermediate Pole (CIP) between the Terrestrial Intermediate Origin (TIO) and the Celestial Intermediate Origin (CIO), positively in the retrograde direction. It is related to UT1 by a conventionally adopted expression in which ERA is a linear function of UT1. Its time derivative is the Earth's angular velocity.

eccentric anomaly: in undisturbed elliptic motion, the angle measured at the center of the ellipse from pericenter to the point on the circumscribing auxiliary circle from which a perpendicular to the major axis would intersect the orbiting body.

eccentricity: a parameter that specifies the shape of a conic section; one of the standard elements used to describe an elliptic orbit.

eclipse: the obscuration of a celestial body caused by its passage through the shadow cast by another body.

eclipse, annular: a solar eclipse in which the solar disk is never completely covered but is seen as an annulus or ring at maximum eclipse. An annular eclipse occurs when the apparent disk of the Moon is smaller than that of the Sun.

eclipse, lunar: an eclipse in which the Moon passes through the shadow cast by the Earth. The eclipse may be total (the Moon passing completely through the Earth's umbra), partial (the Moon passing partially through the Earth's umbra at maximum eclipse), or penumbral (the Moon passing only through the Earth's penumbra).

eclipse, solar: an eclipse in which the Earth passes through the shadow cast by the Moon. It may be total (observer in the Moon's umbra), partial (observer in the Moon's penumbra), or annular.

ecliptic: the plane perpendicular to the mean heliocentric orbital angular momentum vector of the Earth-Moon barycenter in the Barycentric Celestial Reference System, commonly the mean plane of the Earth's orbit.

elements, Besselian: quantities tabulated for the calculation of accurate predictions of an eclipse or occultation for any point on or above the surface of the Earth.

elements, orbital: parameters that specify the position and motion of a body in orbit.

epact: the number of days in the age of the moon on 1 January of any given year in the Gregorian calendar system.

ephemeris: a tabulation of the positions of a celestial object in an orderly sequence for a number of dates.

ephemeris second: the second defined in 1960 as $1/31\,556\,925.9747$ of the tropical year for 1900 January 0 12 hours ET.

ephemeris time (ET): the time scale used from 1960 to 1984 as the independent variable in gravitational theories of the solar system, with its unit and origin conventionally defined. It was superseded by TT and TDB.

epicycle: in representing the motion of planets with epicycles, a circular orbit, whose center moves uniformly over a circular orbit around the Earth called the deferent.

epoch: a fixed instant of time used as a chronological reference datum for calendars, celestial reference systems, star catalogs, or orbital motions.

equation of center: in elliptic motion, the true anomaly minus the mean anomaly. It is the difference between the actual angular position in the elliptic orbit and the position the body would have if its angular motion were uniform.

equation of the equinoxes: the right ascension of the mean equinox referred to the true equator and equinox; alternatively the difference between apparent sidereal time and mean sidereal time ($GAST - GMST$).

equation of the origins: distance between the CIO and the equinox along the intermediate equator; it is the CIO right ascension of the equinox; alternatively

the difference between the Earth Rotation Angle and Greenwich apparent sidereal time (ERA – GAST).

equation of time: the hour angle of the true Sun minus the hour angle of the fictitious mean sun; alternatively, apparent solar time minus mean solar time.

equator: the great circle on the surface of a body formed by the intersection of the surface with the plane passing through the center of the body perpendicular to the reference axis.

equinox: either of the two points on the celestial sphere at which the ecliptic intersects the celestial equator; also, the time at which the Sun passes through either of these intersection points; that is, when the apparent longitude of the Sun is 0° or 180° .

era: a system of chronological notation reckoned from a given date.

escapement: a device that controls the continuous motion of the clock's driving mechanism using the periodic motion of its regulator.

fictitious mean sun: an imaginary body introduced to define mean solar time; essentially the name of a mathematical formula that defined mean solar time. This concept is no longer used in high precision work.

flattening: a parameter that specifies the degree by which the shape of an ellipsoid of revolution differs from a sphere; the ratio $f = (a - b)/a$, where a is the equatorial radius and b is the polar radius.

flicker frequency noise: a type of statistical noise observed in the output frequency of a oscillator that has a $1/f$ spectrum, where f is the sampling frequency.

flicker phase noise: a type of statistical noise observed in the timing signal of a clock that has a $1/f$ spectrum, where f is the sampling frequency.

foliot: in a verge and foliot clock, the crossbar on the verge that controls the time it takes for the verge to rotate.

frame bias: the three offsets of the mean equator and (dynamical) mean equinox of J2000.0, provided by the current model, with respect to the GCRS; the first two offsets are the mean pole offsets at J2000.0 and the third is the offset in right ascension of the mean dynamical equinox of J2000.0.

free core nutation (FCN): free retrograde diurnal mode in the motion of the Earth's rotation axis with respect to the Earth, due to nonalignment of the rotation axis of the core and of the mantle; it is a long-period (of 432 days) free nutation of the CIP in the GCRS.

free falling frame: an isolated local frame which is electrically and magnetically shielded, sufficiently small that inhomogeneities in external fields can be ignored, and self-gravitating effects are negligible.

frequency: the number of cycles or complete alternations per unit time of a carrier wave, band, or oscillation.

frequency comb: a precise tool for measuring different colors, or frequencies, of light in which a mode-locked laser provides an optical spectrum consisting of equidistant lines, which can be used as an optical ruler.

frequency standard: a stable oscillator whose output is used as a precise frequency reference; a primary frequency standard is one whose frequency corresponds to the adopted definition of the second, with its specified accuracy achieved without calibration of the device.

Galileo: a European satellite navigation system.

Gaussian gravitational constant ($k = 0.017\,202\,098\,95$): the gravitational constant specified in units of length (astronomical unit), mass (solar mass) and time (day).

geocentric: with reference to, or pertaining to, the center of the Earth.

Geocentric Celestial Reference System (GCRS): a system of geocentric space-time coordinates within the framework of General Relativity with metric tensor specified by the IAU. The GCRS is defined such that the transformation between BCRS and GCRS spatial coordinates contains no rotation component, so that GCRS is kinematically nonrotating with respect to BCRS. The equations of motion of, for example, an Earth satellite with respect to the GCRS will contain relativistic Coriolis forces that come mainly from geodesic precession. The spatial orientation of the GCRS is derived from that of the BCRS, that is, by the orientation of the ICRS.

geocentric coordinates: the latitude and longitude of a point on the Earth's surface relative to the center of the Earth; also, celestial coordinates given with respect to the center of the Earth.

Geocentric Coordinate Time (TCG): coordinate time of the GCRS based on the SI second. It is related to Terrestrial Time (TT) by a conventional linear transformation.

Geocentric Terrestrial Reference System (GTRS): a system of geocentric space-time coordinates within the framework of General Relativity, co-rotating with the Earth, and related to the GCRS by a spatial rotation which takes into account the Earth orientation parameters. It was adopted by IUGG 2007 Resolution 2. It replaces the previously defined Conventional Terrestrial Reference System.

geodesic precession and nutation: a Coriolis-like effect from relativistic theory in the transformations of the fixed directions of the GCRS referred to the BCRS – the largest components of the relativistic rotation of the GCRS with respect to a dynamically nonrotating geocentric reference system in the framework of General Relativity. Geodesic precession is the secular part of the rotation, and geodesic nutation is the periodic part. Geodesic precession and nutation are included in the IAU 2000 precession-nutation model.

geodetic coordinates: the latitude and longitude of a point on the Earth's surface determined using the geodetic vertical (normal to the specified spheroid).

geoid: an equipotential surface that coincides with mean sea level in the open ocean. On land it is the level surface that would be assumed by water in an imaginary network of frictionless channels connected to the ocean.

geometric position: the geocentric position of an object on the celestial sphere referred to the true equator and equinox, but without including the displacement due to planetary aberration.

getter: component of a Cesium beam tube frequency standard that absorbs Cesium atoms.

GLONASS: a Russian satellite navigation system.

Global Positioning System (GPS): a U.S. satellite navigation system.

GPS Time: a time scale based on the clocks of the GPS that is maintained within better than 1 microsecond of UTC, or TAI, modulo one second. It is ahead of TAI by 19s and differs from UTC depending on the number of leap seconds, since GPS Time is not adjusted for leap seconds.

gravitational constant: denoted G , is an empirical physical constant used to specify the gravitational attraction between objects.

great circle: the circle formed by the intersection of a sphere with a plane that passes through the center of the sphere.

great empirical term: a periodic term introduced into lunar theories prior to the twentieth century in an attempt to make the theories agree with the observations.

Greenwich Apparent Sidereal Time (GAST): the hour angle of the true equinox from the Terrestrial Intermediate Origin (TIO) meridian. Also Greenwich Sidereal Time.

Greenwich Civil Time (GCT): Term used from 1925 until 1952 to refer to Greenwich Mean Time reckoned from midnight, officially replaced by Universal Time in 1952.

Greenwich Mean Astronomical Time (GMAT): the name recommended for use to designate mean solar time reckoned from noon at Greenwich when a 12-hour discontinuity was introduced in 1925 in GMT.

Greenwich mean sidereal time (GMST): Greenwich hour angle of the mean equinox of date defined by a conventional relationship to the Earth Rotation Angle, or equivalently to UT1.

Greenwich Mean Time (GMT): currently the civil time of the United Kingdom. Once the predecessor of UTC for civil time and UT1 for celestial navigation. Prior to 1925 hours counted from noon.

Greenwich sidereal date (GSD): the number of sidereal days elapsed at Greenwich since the beginning of the Greenwich sidereal day that was in progress at Julian date 0.0.

Greenwich sidereal day number: the integral part of the Greenwich sidereal date.

Gregorian calendar: the calendar introduced by Pope Gregory XIII in 1582 to replace the Julian calendar, the calendar now used as the civil calendar in most countries. Every year that is exactly divisible by four is a leap year, except for centennial years, which must be exactly divisible by 400 to be leap years. Thus, 2000 is a leap year, but 1900 and 2100 are not leap years.

height: elevation above ground or distance upwards from a given level (especially sea level) to a fixed point.

heliocentric: with reference to, or pertaining to, the center of the Sun.

heliacal rising: rising of a celestial object at the same time as the Sun rises.

high altitude winds: winds in the upper atmosphere.

horizon: a plane perpendicular to the direction of the zenith. The great circle formed by the intersection of the celestial sphere with a plane perpendicular to the direction of the zenith is called the astronomical horizon.

horizontal parallax: angle subtended by the Earth's radius at the distance of a celestial body.

horologium: historical term for a clock, either a water clock or a verge and foliot clock.

hour angle: angular distance measured westward along the celestial equator from the meridian to the hour circle that passes through a celestial object.

hour circle: a great circle on the celestial sphere that passes through the celestial poles and is therefore perpendicular to the celestial equator.

hydrogen maser: device that produces coherent electromagnetic waves through amplification due to stimulated emission of the hydrogen atom to serve as a precision frequency reference.

hyperfine levels: energy levels in the structure of individual atoms that originate from the interaction of the magnetic moments of the electron and the nucleus.

IAU 2000A precession nutation model: the IAU-recommended precession nutation model representing the motion of the CIP for those who need a model at 0.2 mas level. An abridged model, designated IAU 2000B, is available for those who require a model at the 1 mas level. Standard programs are available from IERS and SOFA.

inclination: the angle between two planes or their poles; usually the angle between an orbital plane and a reference plane; one of the standard orbital elements that specifies the orientation of an orbit.

International Atomic Time (TAI): the continuous scale resulting from analyses by the Bureau International des Poids et Mesures of atomic time standards in many countries. The fundamental unit of TAI is the SI second, and the epoch is 1958 January 1.

International Celestial Reference Frame (ICRF): a set of extragalactic objects whose adopted positions and uncertainties realize the International Celestial Reference System axes and give the uncertainties of the axes. It is also the name of the radio catalog whose 212 defining sources make it currently the most accurate realization of the ICRS. The orientation of the ICRF catalog is within the errors of the standard stellar and dynamic frames at the time of adoption. Successive revisions of the ICRF are intended to minimize rotation from its original orientation. Other realizations of the ICRS have specific names (e.g., Hipparcos Celestial Reference Frame).

International Celestial Reference System (ICRS): the idealized barycentric coordinate system to which celestial positions are referred. It is kinematically nonrotating with respect to the ensemble of distant extragalactic objects. It has no intrinsic orientation but was aligned close to the mean equator and dynamical equinox of J2000.0 for continuity with previous fundamental reference systems. Its orientation is independent of epoch, ecliptic, or equator and is realized by a list of adopted coordinates of extragalactic sources.

International Terrestrial Reference Frame (ITRF): a realization of ITRS by a set of instantaneous coordinates (and velocities) of reference points distributed on the topographic surface of the Earth (mainly space geodetic stations and related markers). Currently the ITRF provides a model for estimating, to high accuracy, the instantaneous positions of these points, which is the sum of conventional corrections provided by the IERS Convention center (solid Earth tides, pole tides) and of a ‘regularized’ position. At present, the latter is modeled by a piecewise linear function, the linear part accounting for such effects as tectonic plate motion, postglacial rebound, and the piecewise aspect representing discontinuities such as seismic displacements. The initial orientation of the ITRF is that of the BIH Terrestrial System at epoch 1984.0.

International Terrestrial Reference System (ITRS): the ITRS is the specific GTRS for which the orientation is operationally maintained in continuity with past international agreements. The co-rotation condition is defined as no residual rotation with regard to the Earth’s surface, and the geocenter is understood as the center of mass of the whole Earth system, including oceans and atmosphere. For continuity with previous terrestrial reference systems, the first alignment was close to the mean equator of 1900 and the Greenwich meridian. The ITRS was adopted as the preferred GTRS for scientific and technical applications and is the recommended system to express positions on the Earth.

ITRF zero-meridian: the plane passing through the geocenter, ITRF pole, and ITRF x -origin.

ITRS CIP coordinates: direction cosines of the CIP in the ITRS, also called pole coordinates. They are currently expressed in the form of x and y coordinates, in arcseconds, the values of which represent the corresponding angles with respect to the polar axis of the ITRS. The sign convention is such that x is positive toward the x -origin of the ITRS and y is in the direction 90° to the west of x .

invariable plane: the plane through the center of mass of the solar system perpendicular to the angular momentum vector of the solar system.

J2000.0: defined in the framework of General Relativity by IAU as being the event (epoch) at the geocenter and at the date 2000 January 1.5 TT = Julian Date 245 1545.0 TT. Note that this event has different dates in different time scales.

Julian calendar: the calendar introduced by Julius Caesar in 46 B.C. to replace the Roman calendar. In the Julian calendar a common year is defined to comprise 365 days, and every fourth year is a leap year comprising 366 days. The Julian calendar was superseded by the Gregorian calendar.

Julian date (JD): the interval of time in days and fraction of a day since 4713 B.C. January 1, Greenwich noon, Julian proleptic calendar. In precise work the timescale, for example, dynamical time or Universal Time, should be specified.

Julian date, modified (MJD): the Julian date minus 2400000.5.

Julian day number: the integral part of the Julian date.

Julian proleptic calendar: the calendar system employing the rules of the Julian calendar, but extended and applied to dates preceding the introduction of the Julian calendar.

Julian year: a period of 365.25 days. This period served as the basis for the Julian calendar.

Laplacian plane: for planets see invariable plane; for a system of satellites, the fixed plane relative to which the vector sum of the disturbing forces has no orthogonal component.

latitude, celestial: angular distance on the celestial sphere measured north or south of the ecliptic along the great circle passing through the poles of the ecliptic and the celestial object.

latitude, terrestrial: angular distance on the Earth measured north or south of the equator along the meridian of a geographic location.

leap second: a second added or subtracted at announced times to keep UTC within 0.9 s of UT1. Generally, leap seconds are applied at the end of June or December.

length of day: strictly the number of fixed length seconds in the day determined from the rotation of the Earth, but most often used to refer to the excess length of day or the difference between the length of day and 86 400 SI seconds.

Lense-Thirring effect: also referred to as frame dragging. The rotation of an object alters the space and time, dragging a nearby object out of its position compared to the predictions of Newtonian physics.

librations: variations in the orientation of the Moon's surface with respect to an observer on the Earth. Physical librations are due to variations in the orientation of the Moon's rotational axis in inertial space. The much larger optical librations are due to variations in the rate of the Moon's orbital motion, the obliquity of the Moon's equator to its orbital plane, and the diurnal changes of geometric perspective of an observer on the Earth's surface.

light-time: the interval of time required for light to travel from a celestial body to the Earth. During this interval the motion of the body in space causes an angular displacement of its apparent place from its geometric place.

light-year: the distance that light traverses in a vacuum during one year.

limb: the apparent edge of the Sun, Moon, a planet or any other celestial body with a detectable disk.

limb correction: correction that must be made to the distance between the center of mass of the Moon and its limb. These corrections are due to the irregular surface of the Moon and are a function of the librations in longitude and latitude and the position angle from the central meridian.

Liouville equation: basic mathematical formulation used to describe the changes in the angular momentum of a rotating system in response to applied torques.

local sidereal time: the local hour angle of equinox.

longitude, celestial: angular distance measured eastward along the ecliptic from the dynamical equinox to the great circle passing through the poles of the ecliptic and the celestial object.

longitude, terrestrial: angular distance measured along the Earth's equator from the Greenwich meridian to the meridian of a geographic location.

Lorentz transformation: mathematical formulation to convert between two different observers' measurements of space and time, where one observer is in constant motion with respect to the other.

lunar phases: cyclically recurring apparent forms of the Moon. New moon, first quarter, full moon and last quarter are defined as the times at which the excess of the apparent celestial longitude of the Moon over that of the Sun is 0° , 90° , 180° , and 270° , respectively.

lunar secular acceleration: gradual speeding up of the Moon's orbital motion due to tidal interaction with the Earth, and related to the slowing of the Earth's rotation.

lunation: the period of time between two consecutive new moons.

lunisolar tidal deceleration: the deceleration of the rotation of the Earth due to tidal interactions. Conservation of angular momentum is maintained by the acceleration of the Moon's orbital motion.

maser: an instrument which uses the monochromatic emission from a narrow band in the spectrum of a suitable molecule or atom to control the frequency of a radio-resonant circuit.

magnitude, stellar: a measure on a logarithmic scale of the brightness of a celestial object considered as a point source.

mean anomaly: in undisturbed elliptic motion, the product of the mean motion of an orbiting body and the interval of time since the body passed the pericenter. The mean anomaly is the angle from the pericenter of a hypothetical body moving with a constant angular speed that is equal to the mean motion.

mean distance: the semimajor axis of an elliptic orbit.

mean elements: elements of an adopted reference orbit that approximates the actual, perturbed orbit. Mean elements may serve as the basis for calculating perturbations.

mean equator and equinox: equator and equinox associated with a celestial pole whose direction is determined only by the precession portion of the precession-nutation transformation.

mean motion: in undisturbed elliptic motion, the constant angular speed required for a body to complete one revolution in an orbit of a specified semimajor axis.

mean place: position of an object on the celestial sphere referred to the mean equator and equinox at a standard epoch.

mean solar time: a measure of time based conceptually on the diurnal motion of the fictitious mean sun, under the assumption that the Earth's rate of rotation is constant.

meridian: a great circle passing through the reference poles and through the zenith of any location on Earth. For planetary observations a meridian is half the great circle passing through the planet's poles and through any location on the planet.

metric space: a set where a notion of distance between elements is defined. The choice of the metric defines the geometric properties of the space.

microwave standard: device providing time or frequency using energy level transitions in atoms with wavelengths ranging from 1 mm to 1 m.

moment of inertia: measure of an object's resistance to changes in rotational velocity.

month: the period of one complete synodic or sidereal revolution of the Moon around the Earth; also, a calendrical unit that approximates the period of revolution of the Moon.

moonrise, moonset: the times at which the apparent upper limb of the Moon is on the astronomical horizon; that is, when the true zenith distance, referred to the center of the Earth, of the central point of the disk is $90^\circ 34' + s - \pi$, where s is the Moon's semi-diameter, π is the horizontal parallax, and $34'$ is the adopted value of horizontal refraction.

nadir: the point on the celestial sphere diametrically opposite to the zenith.

Newcomb's Theory of the Sun: the theory of the Sun developed by Simon Newcomb and published in *The Astronomical Papers for the American Ephemeris*, volume VI, in 1898. It was the basis of the ephemeris of the Sun, time scales, and astronomical constants for about 80 years.

Newtonian mechanics: description of motions based on Newton's Universal Law of Gravity.

node: either of the points on the celestial sphere at which the plane of an orbit intersects a reference plane. The position of a node is one of the standard orbital elements used to specify the orientation of an orbit.

nonrigid Earth: model of the Earth that includes the effects of core flattening, core-mantle interactions, oceans, etc.

nonrotating origin: in the context of the GCRS or the ITRS, the point on the intermediate equator such that its instantaneous motion with respect to the system (GCRS or ITRS as appropriate) has no component along the intermediate equator (i.e., its instantaneous motion is perpendicular to the intermediate equator). It is called the CIO and TIO in the GCRS and ITRS, respectively.

numerical integration: the process of calculating the motion of a body by integrating the equations of motion.

nutation: the short-period oscillations in the motion of the pole of rotation of a freely rotating body that is undergoing torque from external gravitational forces.

obliquity: in general, the angle between the equatorial and orbital planes of a body or, equivalently, between the rotational and orbital poles. For the Earth the obliquity of the ecliptic is the angle between the planes of the equator and the ecliptic.

occultation: the obscuration of one celestial body by another of greater apparent diameter; especially the passage of the Moon in front of a star or planet, or the disappearance of a satellite behind the disk of its primary. If the primary source of illumination of a reflecting body is cut off by the occultation, the phenomenon is also called an eclipse. The occultation of the Sun by the Moon is a solar eclipse.

opposition: a configuration of the Sun, Earth, and a planet in which the apparent geocentric longitude of the planet differs by 180° from the apparent geocentric longitude of the Sun.

optical clock: time standards based on energy level transitions in atoms that provide precise frequencies in the optical range.

optical molasses: a cloud of atoms slowed in three dimensions by laser cooling.

optical pumping: process in which electromagnetic energy is used to raise (or ‘pump’) electrons from a lower energy level in an atom or molecule to a higher one.

orbit: the path in space followed by a celestial body.

orientation: the set of direction angles made by the axes of one coordinate system with the axes of the other.

oscillation: the repeated movement of an object, at equal time intervals, from one point to another and back again.

osculating elements: a set of parameters that specifies the instantaneous position and velocity of a celestial body in its perturbed orbit. Osculating elements describe the unperturbed (two-body) orbit that the body would follow, if perturbations were to cease instantaneously.

parallax: the difference in apparent direction of an object as seen from two different locations; annual parallax refers to the difference in directions as seen from the barycenter and the geocenter, while diurnal parallax refers to the component of parallax due to the observer’s separation from the geocenter.

parsec: the distance at which one astronomical unit subtends an angle of one second of arc; equivalently, the distance to an object having an annual parallax of one second of arc.

Paul trap: a quadrupole ion trap that exists in both linear and three-dimensional varieties that uses constant DC and radio frequency oscillating fields to trap ions.

pendulum clock: clock whose timing is based on the swinging of a pendulum.

Penning trap: a quadrupole ion trap using a constant static magnetic field and spatially inhomogeneous static electric field to trap ions.

penumbra: the portion of a shadow in which light from an extended source is partially but not completely cut off by an intervening body; the area of partial shadow surrounding the umbra.

pericenter: the point in an orbit that is nearest to the center of force.

perigee: the point at which a body in orbit around the Earth most closely approaches the Earth. Perigee is sometimes used with reference to the apparent orbit of the Sun around the Earth.

perihelion: the point at which a body in orbit around the Sun most closely approaches the Sun.

period: the interval of time required to complete one revolution in an orbit or one cycle of a periodic phenomenon.

perturbations: deviations between the actual orbit of a celestial body and an assumed reference orbit; also, the forces that cause deviations between the actual and reference orbits. Perturbations, according to the first meaning, are usually calculated as quantities to be added to the coordinates of the reference orbit to obtain the precise coordinates.

phase: in astronomical phenomena, the ratio of the illuminated area of the apparent disk of a celestial body to the area of the entire apparent disk taken as a circle. For eclipses, phase designations (total, partial, penumbral, etc.) provide general descriptions of the phenomena. More generally, for use with oddly shaped bodies, phase might be defined as $0.5 (1 + \cos(\text{phase angle}))$; the phase of an oscillation or wave is the fraction of a complete cycle corresponding to an offset in the displacement from a specified reference point.

phase angle: the angle measured at the center of an illuminated body between the light source and the observer.

photometry: a measurement of the intensity of light usually specified for a specific frequency range.

piezoelectric effect: the ability of some materials (e.g., quartz crystals) to generate an electric potential in response to applied mechanical stress.

Planck's Constant: (denoted h) a physical constant to describe the sizes of quanta in quantum mechanics. It is the proportionality constant between energy of a photon and the frequency of its associated electromagnetic wave.

polar motion: the motion of the Earth's pole with respect to the ITRS.

post-glacial rebound: the rise of land masses that were depressed by the weight of ice sheets during the last glacial period. Also called, continental rebound, isostatic rebound, post-ice age isostatic recovery.

precession-nutation: the ensemble of effects of external torques on the motion in space of the rotation axis of a freely rotating body or, alternatively, the forced motion of the pole of rotation due to those external torques. In the case of the Earth, a practical definition is that precession-nutation is the motion of the CIP in the GCRS, including the free core nutation and other corrections to the standard models. **Precession** is the secular part of this motion plus the term of 26 000-year period and **nutation** is that part of the CIP motion not classed as precession.

precession of the ecliptic: the secular part of the motion of the ecliptic with respect to the fixed ecliptic.

precession of the equator (and CIP): the uniformly progressing motion of the pole of rotation of a freely rotating body undergoing torque from external gravitational forces. In the case of the Earth, the precession of the equator is caused by solar system objects acting on the Earth's equatorial bulge making the pole of rotation describe a 26 000 year orbit around the ecliptic pole.

precession of the equinox: results from both the precession of the equator and the precession of the ecliptic.

precision: a measure of the tendency of a set of measures to cluster about a value determined by the set.

prograde annual wobble: the component of polar motion of the Earth with a period of one year.

prograde semiannual wobble: the component of polar motion of the Earth with a period of one half a year.

proper motion: the projection onto the celestial sphere of the space motion of a star relative to the solar system; thus, the angular transverse component of the space motion of a star with respect to the solar system. Proper motion is usually tabulated in star catalogs as changes in right ascension and declination per year or century.

proper time: time measured along the trajectory of an observer in space time and invariant in any coordinate change.

pulsar: highly magnetized rotating neutron star that is observed to emit periodically a beam of radio waves with a period ranging between 1.5 ms and 8.5 s.

quality factor (Q): the quality factor for oscillating systems is defined as the ratio of total energy in a system to the energy lost per cycle; also the ratio that provides an indication of the sharpness of the peak of a resonance curve.

quantum logic clock: a frequency standard based on optical frequencies produced by energy level transitions in single ions.

quartz crystal clock: a clock based on the use of quartz crystal to regulate the oscillations in an electrical circuit.

radial velocity: the rate of change of the distance to an object.

Ramsey fringe: the sinusoidal fringe pattern in the excited-state population, as a function of detuning from resonance, caused by the interaction between an atomic wave function and a probe, which is transferred between the ground and excited states, depending upon the phase between the excitation field and the atomic oscillation.

random walk frequency noise: noise in a signal inversely proportional to the sampling frequency squared.

refraction, astronomical: the bending of a ray of light as it passes through the Earth's atmosphere. Most commonly calculated using pressure, temperature, humidity and wavelength.

residuals: the difference between actual observations and theoretical estimates of a variable.

retrograde motion: for orbital motion in the solar system, motion that is clockwise in the orbit as seen from the north pole of the ecliptic; for an object observed on the celestial sphere, motion that is from east to west, resulting from the relative motion of the object and the Earth.

Riemannian space time: the four-dimensional coordinates of space and time in Riemannian geometry, which is the study of smooth curved surfaces. This is compared to Euclidean geometry, which is the study of flat spaces.

right ascension: angular distance given either in arc or time units measured eastward along the celestial equator from the equinox, or CIO, to the hour circle passing through the celestial object, usually given in combination with declination.

Sagnac effect: the interference observed in two beams of light sent in opposite directions around a rotating loop, caused by the fact that the pulse sent in the same direction as the rotation of the loop must travel further than the pulse sent in the opposite direction of the rotation.

seasonal hours: hours of variable length due to the seasonal change in the length of daylight.

second, Système International (SI): the duration of 9 192 631 770 cycles of radiation corresponding to the transition between two hyperfine levels of the ground state of Cesium 133.

secular acceleration of the Moon: the acceleration of the mean motion of the Moon due to gravitational perturbations and tidal retardation of Earth rotation.

secular deceleration of Earth: change in the length of the day by about 1.7 ms per day per century, due to a combination of lunar tides and post-glacial rebound of portions of the Earth's crust.

secular polar motion: a nonperiodic motion of the Earth's pole toward the direction of approximately 75° west longitude.

selenocentric: with reference to, or pertaining to, the center of the Moon.

semidiameter: the angle at the observer subtended by the equatorial radius of the Sun, Moon, or a planet.

semimajor axis: half the length of the major axis of an ellipse; a standard element used to describe an elliptical orbit.

Shortt clock: a pendulum clock in which a ‘slave pendulum’ performs the mechanical work such as turning the hands of the clock, opening and closing contacts, and giving impulses to the master pendulum, thus leaving the master pendulum free of most perturbations.

sidereal day: the interval of time between two consecutive transits of the equinox.

sidereal hour angle: angular distance measured westward along the celestial equator from the catalog equinox to the hour circle passing through the celestial object. It is equal to 360° minus the right ascension in degrees.

sidereal time: the measure of time defined by the apparent diurnal motion of the equinox; hence, a measure of the rotation of the Earth with respect to the stars rather than the Sun.

signal-to-noise ratio: the ratio between the amplitude of the signal to that of the noise.

Sisyphus cooling: a mechanism for laser cooling of atoms using light forces. The mechanism involves a polarization gradient that introduces nonconservative light forces, which can reduce the average kinetic energy of atoms.

solar time: time based on the rotation of the Earth with respect to the position of the Sun.

solstice: either of the two points on the ecliptic at which the apparent longitude of the Sun is 90° or 270° ; also, the time at which the Sun is at either point.

spacetime metric: a chosen specification defining the properties of time and distance in the general theory of relativity.

stability: ability of a standard to maintain its synchronization, or syntonization, over time. A stable clock would produce the same measures over a range of time intervals.

standard epoch: a date and time used to specify a reference frame. Prior to 1984 coordinates of star catalogs were commonly referred to the mean equator and equinox of the beginning of a Besselian year. Beginning with 1984 the Julian year has been used, as denoted by the prefix J, for example, J2000.0.

sunrise, sunset: the times at which the apparent upper limb of the Sun is on the astronomical horizon; that is, when the true zenith distance, referred to the center of the Earth, of the central point of the disk is $90^\circ 50'$, based on adopted values of $34'$ for horizontal refraction and $16'$ for the Sun’s semidiameter.

synchronization: the process of setting two standards to read the same time.

synodic period: time taken for an object to reappear at the same point in the sky, relative to the Sun, as observed from Earth.

syntonization: the process of setting two standards to the same frequency.

Teph: the independent time argument of the JPL and MIT/CfA solar-system ephemerides. It differs from Barycentric Coordinate Time (TCB) by an offset and a constant rate. The linear drift between Teph and TCB is such that the rates of Teph and TT are as close as possible for the time span covered by the particular ephemeris. Each ephemeris defines its own version of Teph; the Teph of the JPL ephemeris DE405 is for practical purposes the same as TDB.

Terrestrial Dynamical Time (TDT): time scale for apparent geocentric ephemerides defined by a 1979 IAU resolution and replaced by Terrestrial Time (TT) in 1991.

Terrestrial Ephemeris Origin (TEO): the original name for the Terrestrial Intermediate Origin (TIO) given in the IAU 2000 resolutions.

Terrestrial Intermediate Origin (TIO): origin of longitude in the Intermediate Terrestrial Reference System. It is the nonrotating origin in the ITRS that is recommended by the IAU, where it was originally designated Terrestrial Ephemeris Origin. The name Terrestrial Intermediate Origin was adopted by IAU in 2006. The TIO was originally set at the ITRF origin of longitude, and throughout 1900–2100 it stays within 0.1 mas of the ITRF zero meridian.

Terrestrial Intermediate Reference System (TIRS): a geocentric reference system defined by the intermediate equator of the CIP and the TIO. It is related to the ITRS by polar motion and s' (TIO locator). It is related to the Celestial Intermediate Reference System by a rotation of ERA around the CIP, which defines the common z -axis of the two systems.

Terrestrial Time (TT): a coordinate time whose mean rate is close to the mean rate of the proper time of an observer located on the rotating geoid. At 1977 January 1.0 TAI exactly, the value of TT was 1977 January 1.000 372 5 exactly. It is related to the Geocentric Coordinate Time (TCG) by a conventional linear transformation. An accurate realization of TT is TT (TAI) = TAI + 32 s.184. In the past TT was called Terrestrial Dynamical Time (TDT).

terminator: the boundary between the illuminated and dark areas of the apparent disk of the Moon, a planet, or a planetary satellite.

tidal acceleration, tidal deceleration: the rate of slowing of the Earth's rotation caused by tidal forces between the Moon and the Earth causing a gradual increase of the distance between the Earth and the Moon.

tidal friction: the frictional force caused by the interaction between the tides and the Earth's surface.

tides, ocean: periodic rise and fall of ocean waters due to the attraction of the Moon and Sun.

tides, solid Earth: periodic rise and fall of areas of the Earth crust due to the attraction of the Moon and Sun.

time dilation: phenomena where an observer finds that another clock, physically identical to the observer's clock, is ticking at a slower rate as measured by his own clock. In special relativity, clocks moving with respect to an inertial system are measured to be running slower. In general relativity, clocks at lower potentials in a gravitational field are measured to be running slower.

time variation (TVAR): a statistic used to describe clock noise.

TIO locator (denoted s'): the difference between the ITRS longitude and the instantaneous longitude of the intersection of the ITRS and intermediate equators. The TIO was originally set at the ITRF origin of longitude. As a consequence of polar motion the TIO moves according to the kinematical property of the nonrotating origin. The TIO is currently located using the quantity s' , the rate of which is of the order of $50\mu\text{ascy}^{-1}$ and is due to the current polar motion.

TIO meridian: moving plane passing through the geocenter, the CIP, and the TIO.

topocentric: with reference to, or pertaining to, a point on the surface of the Earth.

transit: the passage of the apparent center of the disk of a celestial object across a meridian; also, the passage of one celestial body in front of another of greater apparent diameter (e.g., the passage of Mercury or Venus across the Sun or Jupiter's satellites across its disk); however, the passage of the Moon in front of the larger apparent Sun is called an annular eclipse. The passage of a body's shadow across another body is called a shadow transit; however, the passage of the Moon's shadow across the Earth is called a solar eclipse.

trapped ions: ions trapped in an electromagnetic field to provide a timing source.

true anomaly: the angle, measured at the focus nearest the pericenter of an elliptical orbit, between the direction of the pericenter and the radius vector from the focus to the orbiting body; one of the standard orbital elements.

true equator and equinox: the celestial coordinate system determined by the instantaneous positions of the celestial equator defined by the Celestial Intermediate Pole and the ecliptic.

twilight: the interval of time preceding sunrise and following sunset during which the sky is partially illuminated. Civil twilight comprises the interval when the zenith distance, referred to the center of the Earth, of the central point of the Sun's disk is between $90^\circ 50'$ and 96° ; nautical twilight comprises the interval from 96° to 102° ; astronomical twilight comprises the interval from 102° to 108° .

umbra: the portion of a shadow cone in which none of the light from an extended light source (ignoring refraction) can be observed.

uniform time scale: a time scale having the same unit of time consistently.

Universal Time (UT): a measure of time that conforms, within a close approximation, to the mean diurnal motion of the Sun and serves as the basis of all civil timekeeping. The term 'UT' is used to designate a member of the family of Universal Time scales (e.g., UTC, UT1).

Universal Time (UT1): angle of the Earth's rotation about the CIP axis defined by its conventional linear relation to the Earth Rotation Angle (ERA). It is related to Greenwich apparent sidereal time through the ERA, and determined by observations. UT1 can be regarded as a time determined by the rotation of the Earth. It can be obtained from the uniform time scale UTC by using the quantity $UT1 - UTC$.

UT1 – UTC: difference between the UT1 parameter derived from observations and the uniform time scale UTC, the latter being currently defined as: $UTC = TAI + n$, where n is an integral number of seconds, such that $|UT1 - UTC| < 0.9s$.

verge: in a verge and foliot clock, a lever, with projections, which intermittently lock the escape wheel and transmits impulses from the escape wheel to the pendulum.

vernal equinox: the ascending node of the ecliptic on the celestial sphere; also, the time at which the apparent longitude of the Sun is 0° .

vertical: the apparent direction of gravity at the point of observation (normal to the plane of a free level surface).

week: an arbitrary period of days, usually seven days; approximately equal to the number of days counted between the four phases of the Moon.

white frequency noise: a type of statistical noise observed in the output frequency of an oscillator that has no dependency on sampling frequency.

white phase noise: a type of statistical noise observed in the timing signal of a clock that has no dependency on sampling frequency.

year: a period of time based on the revolution of the Earth around the Sun. The **calendar year** is the time between two dates with the same designation in a calendar. The **tropical year** is the period of one complete revolution of the mean longitude of the Sun through 360° . The **anomalistic year** is the mean interval between successive passages of the Earth through perihelion. The **sidereal year** is the mean period of revolution with respect to the background stars.

year, Besselian: the period of one complete revolution in right ascension of the fictitious mean sun, as defined by Newcomb. The beginning of a Besselian year, traditionally used as a standard epoch, is denoted by the suffix '.0'. Since 1984 standard epochs have been defined by the Julian year rather than the Besselian year. For distinction, the beginning of the Besselian year is now identified by the prefix B (e.g., B1950.0).

Zeeman effect: the splitting of a spectral line into several components in the presence of a static magnetic field.

zenith: in general, the point directly overhead on the celestial sphere. The astronomical zenith is the extension to infinity of a plumb line. The geocentric zenith is defined by the line from the center of the Earth through the observer. The geodetic zenith is the normal to the geodetic ellipsoid at the observer's location.

zenith distance: angular distance measured along the great circle from the zenith to the celestial object. Zenith distance is 90° minus altitude.

zij: a generic name for an Arabic astronomical book that includes tabular parameters for calculating positions of astronomical bodies.

Index

a

- A.1 200, 201
- A3 201, 227
- AAM, *see* Atmospheric Angular Momentum
- aberration 12, 13, 90
- accuracy 143, 302, 304
- Adams, John Couch 28, 42
- administrative regulations 287
- advanced time 20
- Agenzia Spaziale Italiana 273
- air traffic 305
- Ajisai 273
- d'Alembert, Jean le Rond 30
- Alfonsine Tables 25
- Alfonso X 25
- algorithm 206, 208, 215
- alkali atoms 155, 159, 172
- alkali metals 152, 153, 307
- Allan deviation 147, 148, 156
- Allan variance 147, 148, 200
- Almagest 24, 25, 189
- almanacs 31, 79
- aluminum 184, 185
- American National Standards Institute (ANSI) 299
- American Nautical Almanac 10, 17
- ammonia molecule 157
- ampere 196
- analemma 11
- Analytical Engine 32
- anchor escapement 136, 190
- angular momentum 44, 52, 69, 108, 276
 - atmospheric (AAM) 54, 276, 278, 279
 - oceanic (OAM) 54, 278, 279
- annual component 7
- annual motion 67
- annual variation 51
- ANSI, *see* American National Standards Institute
- Apollo 34
- Apollonius 24
- applications, future 305
- Areocentric Coordinate Time (TCA) 247
- Argoli, Andrea 27
- Aristarchus of Samos 24, 41
- Aristotle 24
- d'Arrest, Heinrich Louis 28
- artificial satellite 36
- Āryabhata 25
- Āryabhatīya 25
- asterisms 130
- astrodynamics 36
- astrolabe 5, 75, 76, 84
- astrometry 37
- Astronomia Carolina 27
- astronomical almanacs 90, 225
- astronomical constants 27, 32, 35, 36, 117
- astronomical observations 75, 85
- astronomical regulators 137
- astronomical unit 36
- atmosphere 276
 - moment of inertia 53
- Atmospheric Angular Momentum (AAM) 54, 276, 278, 279
- atmospheric drag 272, 273
- atomic energy level transition 152, 307, 308
- atomic frequency 82
- atomic physics 236, 247
- atomic resonance frequency 171
- atomic resonator 154, 156, 157
- atomic standards 83
- atomic time 236
- atomic timekeeping 285
- Atomichron 162, 199
- axis 72
 - fundamental 61

- of inertia 71
- polar 61, 66
- rotation 7, 71

b

- Babbage, Charles 32
- banking and finance 301
- barium 177
- barycenter 35, 61, 63, 91, 113–115, 118, 121, 124
- Barycentric Celestial Reference System (BCRS) 35, 63, 106, 109, 119
- Barycentric Coordinate Time (TCB) 105, 109–111, 118–124, 126, 236, 246
- Barycentric Dynamical Time (TDB) 92, 111, 114, 115, 117, 119, 122–125
- Barycentric Ephemeris Time 122
- BCRS, *see* Barycentric Celestial Reference System
- Beidou/Compass 244, 259
- Berthoud, Ferdinand 139
- beryllium 177, 185
- Bessel, Friedrich Wilhelm 3, 67
- Besselian solar year 19
- BIH, *see* Bureau International de l'Heure
- BIPM, *see* Bureau International des Poids et Mesures
- black body radiation 156, 196, 215, 228
- black holes 265
- Bradley, James 3, 30, 66
- Brahe, Tycho 26, 190
- British Nautical Almanac 17, 28
- Brouwer, Dirk 53
- Brown's lunar tables 31, 50
- Brown's lunar theory 86
- Brown's theory 32
- Brown, Ernest William 31, 45, 48
- Burattini, Tito Livio 191
- Bureau International de l'Heure (BIH) 16, 76, 124, 201, 202, 204, 205, 208, 225–229
- Bureau International des Poids et Mesures (BIPM) 120, 202, 204, 207, 215, 217, 218, 220, 221, 228, 281–283, 291
- Circular T 216
- Bürgi, Jost 190

c

- C-field 155
- cable delays 251
- calcium 184
- calendar 239
- Egyptian 130
- Gregorian 2

- calibration 160, 181, 192, 213, 249, 251, 254
- candela 197
- carrier frequency 259
- carrier phase 271
- navigation satellite 259
- catalogs
- FK 19
- FK3 19
- FK4 19, 86
- FK5 5, 19, 64, 86
- fundamental 19
- Hipparcos 19
- NFK 19
- cavity pulling 157
- cavity resonance frequency 172
- CCDS, *see* Comité Consultatif pour la Définition de la Seconde
- CCIF, *see* International Telephone Consultative Committee
- CCIR, *see* International Radio Consultative Committee
- CCIT, *see* International Telegraph Consultative Committee
- CCITT, *see* International Telephone and Telegraph Consultative Committee
- CCTF, *see* Consultative Committee for Time and Frequency
- CDMA, *see* Code Division Multiple Access
- Celestial Ephemeris Pole 64
- Celestial Intermediate Origin (CIO) 15, 63, 64, 68, 69, 73
- Celestial Intermediate Pole (CIP) 15, 63, 64, 66, 68, 70, 71, 73
- Celestial Intermediate Reference System (CIRS) 63, 64
- celestial mechanics 81, 91, 108
- celestial pole offsets 66, 70, 118, 264, 271
- Celestial Reference System 63
- center of mass 270
- Centre Nationale d'Études Spatiales (CNES) 274, 295
- cesium 98, 111, 118, 120, 123, 151, 153, 155, 159, 162, 171–173, 192, 193, 195, 196, 199–201, 226, 228, 241, 257
- cesium beam 159, 299, 307
- standard 200
- tube 162, 163, 165, 169
- cesium fountain 166, 168, 169
- cesium frequency 193
- CGPM, *see* Conférence Général des Poids et Mesures or General Conference on Weights and Measures
- Chandler component 7

- Chandler motion 71
- Chandler wobble 67
- Chandler, Seth Carlo 7, 67, 71
- China Satellite Navigation Project Center (CSNPC) 259
- chip-scale atomic clock 307
- chronometer 138, 139
 - Harrison 5
- CIO, *see* Celestial Intermediate Origin
- CIP 75
- CIP, *see* Celestial Intermediate Pole or Conventional Intermediate Pole
- CIPM, *see* Comité International des Poids et Mesures or International Committee for Weights and Measures
- Circular T 216, 218
- CIRS, *see* Celestial Intermediate Reference System
- Clairaut, Alexis 30
- Cleaton, C. E. 157
- Clemence, Gerald Maurice 81, 82
- Clement, William 136
- clepsydra 130, 131
- clock 37, 101, 102, 114, 119, 124, 129, 130, 192, 205, 208, 213–215, 243, 253, 259
 - atomic 34, 57, 87, 109, 111, 118, 151–154, 156–158, 174, 177, 199, 201, 202, 220, 226, 240, 248
 - atomic beam 158
 - cesium 120
 - cesium beam 161
 - comparison 111, 251
 - crystal-controlled 56
 - distributed 309
 - future 307
 - longcase 136, 137
 - mechanical 6, 10, 57, 132
 - optical 182–184, 307
 - optical ion 183
 - optical neutral atom 184
 - pendulum 5, 51, 79, 133–138, 141, 190
 - performance 142
 - portable 208, 213
 - quantum logic 184
 - quartz 51, 141, 193
 - quartz crystal 57, 79, 139, 148, 199, 226
 - Riefler 137, 138
 - rubidium 172, 201, 244
 - sand 5
 - Shortt 139
 - single ion 184
 - stability 146, 177
 - stored ion 174
 - synchronization 111
 - transition 153, 155
 - transition frequency 176, 184
 - transport 252, 253
 - water 5, 132
 - wheeled 191
 - world line 252
- CNES, *see* Centre Nationale d'Études Spatiales
- coaxial cable 251, 254
- Code Division Multiple Access (CDMA) 299
- comb 184, 185
- Comité Consultatif pour la Définition de la Seconde (CCDS) 83, 87, 120, 191, 202, 204–206, 208, 215, 218, 219
- Comité International des Poids et Mesures (CIPM) 80, 83, 182, 191, 202–205, 219, 281
- Commission Nationale de l'Heure 201
- communications 308
- communications satellite 259
- computers 32
- Comrie, L. J. 32
- Conférence Général des Poids et Mesures (CGPM) 83, 191, 195, 196, 202–204, 218, 219, 228, 240, 281
- Connaissance des Temps 27, 28
- Consultative Committee for Time and Frequency (CCTF) 196, 282
- Consultative Committees 282
- Conventional Intermediate Pole (CIP) 263, 264
- coordinate system
 - Earth-fixed 252, 253
 - geocentric 219
 - space time 235
- Coordinated Universal Time (UTC) 16, 17, 76, 113, 115, 123, 207, 216, 223, 227–229, 231, 232, 245, 246, 255, 257, 258, 308
- Copernican theory 3
- Copernicus, Nicolas 25, 41
- Coriolis force 108
- correlator 266
- cosmic time 224
- cosmopolitan time 223
- Coster, Salomon 134, 136
- Cowell, P. H. 33
- crystal 151
- CSNPC, *see* China Satellite Navigation Project Center
- Curie, Jaques 140
- Curie, Pierre 140
- cycloid 135

d

- Danjon, André 80, 83
- Danjon astrolabes 56
- Darwin, George Howard 52
- day 2, 31
 - length 9
 - mean solar 86
- daylight saving time 20
- de Sitter, Willem 49, 50, 53, 56
- dead-bead escapement 137
- decadal irregularities 54
- decadal polar motion 70
- decadal variations 65, 70
- decans 130
- declination 3
- deferent 24
- deglaciation 65
- Delaunay, Charles-Eugène 30, 42, 43
- delay
 - atmospheric 269
 - group 268
 - ionospheric 250, 269
 - phase 268
 - tropospheric 250
- ΔT 124, 231
- dissemination methods
 - one-way 250
 - two-way 250
- Doppler broadening 172
- Doppler cooling 166, 167
- Doppler effect 102, 156, 246, 274, 276
- Doppler measurements 34
- Doppler observations 75, 264
- Doppler Orbit Determination and Radiopositioning Integrated on Satellite (DORIS) 274, 277, 295
- Doppler, Christian 246
- DORIS, *see* Doppler Orbit Determination and Radiopositioning Integrated on Satellite
- Dudley Observatory 32
- Dunthorne, Richard 42
- DUT1 228, 231, 255

e

- EAL, *see* Échelle Atomique Libre
- Earth 34, 43, 67, 81, 106–108, 110, 111, 118, 119, 121, 123, 124, 213, 219, 237, 241, 242, 244, 247, 248
 - albedo 272, 273
 - axis 70
 - center of mass 245
 - core 54, 63, 89
 - figure 89
 - flattening 32
 - interior 306
 - magnetic field 28
 - mantle 54, 89
 - motion 29
 - orbit 18, 30, 31, 42, 122
 - orbital motion 35, 236, 239
 - orbital velocity 96
 - orientation 57, 61, 69, 75, 76, 237–240, 263–266, 271, 273, 274, 276, 279, 290, 306
 - orientation parameters 64, 264–266, 270, 274, 279, 285, 291, 294, 295
 - rotation 7, 29, 31, 41, 44, 45, 47, 49, 50, 53, 55, 57, 61, 62, 76, 79, 84, 86, 87, 90, 91, 113, 137, 193, 194, 201, 205, 227–229, 231, 236, 239, 240, 252, 253, 263, 270, 285, 288, 290, 294, 306–308
 - rotation angle 14, 61, 68, 69, 84, 85, 229, 264, 279
 - rotation axis 53, 226
 - rotational motion 253
 - rotational speed 52, 63, 65, 81, 119, 160, 226, 232, 246
 - satellites 57
 - Earth Rotation Angle (ERA) 15, 64, 68, 73
 - eccentric circle 30
 - eccentricity 31
 - Échelle Atomique Libre (EAL) 207–209, 213–216
 - steering 215
 - Eckert, Wallace 32
 - eclipse 237, 238
 - lunar 237
 - observation 42
 - solar 84, 237, 239
 - ecliptic 2, 3, 18, 61, 73
 - Einstein's theory of relativity 237
 - Einstein, Albert 29, 92, 103, 104
 - Ekphantus 41
 - electrical power 300
 - electromagnetic coupling 53
 - electromagnetic field 152
 - electromagnetic signals 253
 - electromagnetic spectrum 181
 - electron 151–153, 155
 - e-LORAN 298
 - emergency services 301
 - empirical terms 46
 - ENVISAT 276, 295
 - ephemerides 19, 25, 27, 29, 31, 33, 34, 36, 37, 45, 79, 86, 90, 91, 113, 114, 117, 122, 123, 236–240, 247, 248, 308

- geocentric 218
- lunar 35
- satellite 259
- solar 35
- solar system 35, 57
- ephemeris 85, 121
 - DE 405 43, 123
 - DE 421 43
 - geocentric 115
 - lunar 32, 88
- Ephemeris Time (ET) 35, 37, 57, 79–83, 85, 90, 91, 113, 115, 120, 124, 125, 160, 161, 191–193, 195, 236
- Ephemeris Time Revised (ET_R) 124
- epicycle 24, 30, 237
- equation of time 10, 11, 239
- equator 3, 64, 69, 118
- equinox 3, 12, 31, 34, 45, 56, 61, 64, 69, 85
 - catalog 19, 35, 86
 - dynamical 35, 86, 118
- equivalence principle 34, 104
- ERA, *see* Earth Rotation Angle
- eras 2
- error
 - Type A 213
 - Type B 213
- Essen, Louis 16, 151, 152, 158, 161, 192, 228
- ET_R, *see* Ephemeris Time Revised
- ET, *see* Ephemeris Time
- Etalon-I 273
- Euler, Leonhard 7, 30, 67
- extragalactic objects 118

f

- FAGS, *see* Federation of Astronomical and Geophysical Data Analysis Services
- FDMA, *see* Frequency Division Multiple Access
- Federation of Astronomical and Geophysical Data Analysis Services (FAGS) 289, 290, 294
- Ferrel, William 43
- fictitious mean Sun 9, 10, 14, 15, 19
- field
 - gravitational 285, 306
 - magnetic 285, 306
- FK4 86
- FK5 5, 64, 86
- Flamsteed, John 27, 30, 41
- Fleming, Sandford 223
- foliot regulator 132
- Fotheringham, John K. 48

- Foucault, Léon 41
- fractional frequency 142, 156
- frame synchronization 299
- free atomic time scale 208
- free core nutation 66
- free wobble 63
- frequency 102, 109, 142, 145, 147, 148, 151, 156, 175, 183, 195, 246
 - cesium 160
 - dissemination 250, 254
 - stability 142
 - transfer 249, 258
- frequency comb 181, 182
- Frequency Division Multiple Access (FDMA) 299
- frequency standard 202, 203, 205, 215
 - atomic 199
- Fromanteel, Johannes 136
- fundamental star catalog 32

g

- GA, *see* Greenwich Atomic
- Galilean principle of relativity 95
- Galilei, Galileo 5, 41, 133, 244
- GALILEO 259
- Galle, Johannes Gottfried 28
- Gauss, Carl Friedrich 28
- Gaussian Gravitational Constant 28
- GBR 193
- GCRS, *see* Geocentric Celestial Reference System
- GCT, *see* Greenwich Civil Time
- GEE 243
- General Conference on Weights and Measures (CGPM) 282, 283
- general relativity 95, 98, 103–105, 115, 118, 119, 205, 294
- geocenter 91, 113, 118, 121, 123
- Geocentric Celestial Reference System (GCRS) 15, 63, 64, 66, 68, 70, 106, 108, 109, 119, 251
- Geocentric Coordinate Time (TCG) 106, 110, 111, 118–124, 126, 218–220, 247, 251
- Geocentric Terrestrial Reference System 251
- geodesic 104
- geodesic precession and nutation 63
- geodetic precession 108
- geoid 91, 120, 204, 215, 218, 245, 246, 252, 253
- geomagnetic storms 251
- geophysical modeling 276
- geopotential 272, 273

geostationary orbits 253
 geostationary satellites 251
 Gerard of Cremona 189
 Gill, G.W. 31
 Ginzel, Friedrich 44
 Global Navigation Satellite Systems
 (GNSS) 244, 271, 294, 297, 301, 305,
 308
 Global Positioning System (GPS) 16, 37,
 75, 76, 151, 208, 232, 244, 258, 270–273,
 294, 297, 299, 301
 – System Time 244, 245, 258
 GLONASS 244, 258, 294
 Glydén, J.A.H. 67
 GMAT, *see* Greenwich Mean Astronomical
 Time
 GMST, *see* Greenwich Mean Sidereal Time
 GMT, *see* Greenwich Mean Time
 GNSS, *see* Global Navigation Satellite
 Systems
 Gould, B.A. 32
 GPS, *see* Global Positioning System
 gravitation 27, 41
 gravitational constants 36
 gravitational field 270
 gravitational potential 104, 121
 gravitational red shift 245
 gravitational torques 67
 gravity 103
 gravity field 245, 274
 gravity potential 213, 215
 great empirical term 5, 31, 50, 86
 Greenwich 48
 – apparent solar time 9
 – hour angle 14
 – meridian 9, 10, 13, 17, 62, 64, 81, 223,
 224
 – midnight 10
 Greenwich Atomic (GA) 199
 Greenwich Civil Time (GCT) 10, 17, 225,
 226
 Greenwich Mean Astronomical Time
 (GMAT) 10, 17, 225
 Greenwich Mean Sidereal Time (GMST)
 15, 64
 Greenwich Mean Solar Time 226
 Greenwich Mean Time (GMT) 10, 17,
 223–227
 Greenwich Observatory 202
 Greenwich Sidereal Time (GST) 64, 68, 69
 Greenwich Time 47
 GRGS, *see* Groupe de Recherche de
 Géodésie Spatiale
 ground conductivity 251

Groupe de Recherche de Géodésie Spatiale
 (GRGS) 274
 GST, *see* Greenwich Sidereal Time
 Günther, Wolfgang 134

h

Hadamard variance 148
 Hall, R.G. 192
 Halley, Edmond 3, 27, 28, 30, 41
 Hansen's tables 31
 Hansen, Peter 30
 Harrison, John 138
 Harvard University 200
 Hecker, Johann 27
 Heisenberg uncertainty principle 153
 heliacal rising 130
 Henderson, Thomas 3
 Herschel, William 3, 28
 Hevelius, Johannes 134, 135
 Hiketas of Syracuse 41
 Hipparchus 3, 24, 29, 30, 65
 Hipparchos 118
 Hipparchos Catalog 62
 Hooke, Robert 136
 horologium 129
 hour 189
 – equinoctial 5
 – seasonal 5, 130
 Hupsicles 189
 Huygens, Christiaan 5, 134–136
 hydrogen 153, 158
 hydrogen maser 169, 170, 174, 266, 299
 – active 170, 171
 – cryogenic 170
 – passive 171
 hydrogen-like atoms 157
 hydrology 54
 hyperbolic navigation system 298
 hyperfine clock transition 172
 hyperfine levels 153, 159
 hyperfine states 162
 hyperfine structure 173, 174
 hyperfine transition 155, 163, 193
 – frequency 176
 – of the cesium atom 308

i

IAG, *see* International Association of
 Geodesy
 IAU 1976 System of Astronomical
 Constants 15, 61
 IAU 1980 Theory of Nutation 15, 61, 64
 IAU 2000A precession-nutation model 64,
 66, 70, 118

- IAU 2006 Theory of Precession 70, 118
- IAU resolutions 105
- IAU, *see* International Astronomical Union
- IBM Selective Sequence Electronic Calculator 32
- ibu al-Shatir 25
- ICRF, *see* International Celestial Reference Frame
- ICRS, *see* International Celestial Reference System
- ICSU, *see* International Council for Science
- IDS, *see* International DORIS Service
- IERS
 - Bulletin B 279
 - Rapid Service/Prediction Service 279
- IERS Combination Research Centers 292
- IERS Conventions 61
- IERS, *see* International Earth Rotation and Reference Systems Service *or* International Earth Rotation Service
- IGN, *see* Institut Géographique National
- IGS, *see* International GNSS Service *or* International GPS Service
- ILE, *see* Improved Lunar Ephemeris
- ILRS, *see* International Laser Ranging Service
- ILS, *see* International Latitude Service
- Improved Lunar Ephemeris (ILE) 31, 32, 86
- inertial frame 103
- Innes, R.T.A. 46
- Institut Géographique National (IGN) 292, 295
- intelligent transportation system 298
- interferometry
 - connected element 75
 - very long baseline 75
- international activities 281
- International Association of Geodesy (IAG) 224, 285, 286, 293–295
- International Astronomical Union (IAU) 17, 36, 62, 73, 80, 83, 87, 92, 95, 105, 108, 109, 113, 115, 122, 123, 161, 195, 202, 207, 218, 220, 226–228, 232, 247, 251, 284, 289, 290
- International Atomic Time (TAI) 87, 111, 113–115, 120, 123–125, 199, 201–205, 207, 208, 212, 215–221, 227–229, 258, 308
 - distribution 216
 - formation 217
 - stability 216
- International Celestial Reference Frame (ICRF) 62, 63, 118, 265, 285, 290, 292
- International Celestial Reference System (ICRS) 19, 35, 62–64, 118, 285, 290, 291
 - prime meridian 64
- International Committee for Weights and Measures (CIPM) 16, 161, 202, 228, 282, 283
- International Council for Science (ICS) 284
- International Council for Science (ICSU) 289, 290
- International DORIS Service (IDS) 276, 285, 290, 295
- International Earth Rotation and Reference Systems Service (IERS) 17, 19, 36, 61, 63, 64, 66, 75, 76, 125, 229, 231, 274, 279, 285, 289–291, 293, 295
- International Earth Rotation Service (IERS) 76, 290
- International Geographical Congress 223, 224
- International GNSS (Global Navigational Satellite Service) 294
- International GNSS Service (IGS) 271, 272, 290
- International GPS Service (IGS) 285
- International Laser Ranging Service (ILRS) 274, 275, 285, 290, 293, 294
- International Latitude Service (ILS) 7, 76, 229
- International Meridian Conference 12, 17, 224, 225
- International Polar Motion Service (IPMS) 76
- International Radio Consultative Committee (CCIR) 16, 17, 202, 219, 226–228, 287
 - Study Group 7 16
- International Radiotelegraph Convention 287
- International Research Council 226
- International Scientific Radio Union 202
- international standards 308
- International System of Units (SI) 119, 241
- International Telecommunications Convention 287
- International Telecommunications Regulations 287
- International Telecommunications Union (ITU) 202, 255, 286–289
 - Radiocommunications Sector (ITU-R) 229, 232, 249, 287
 - Telecommunication Standardization Sector (ITU-T) 287
 - Telecommunications Development Sector (ITU-D) 287
- International Telegraph Consultative Committee (CCIT) 287

- International Telegraph Convention 287
 - International Telegraph Union 287
 - International Telephone and Telegraph Consultative Committee (CCITT) 287
 - International Telephone Consultative Committee (CCIF) 287
 - International Terrestrial Reference Frame (ITRF) 63, 64, 285, 290, 292, 294, 295
 - International Terrestrial Reference System (ITRS) 15, 19, 63, 64, 68, 285, 290
 - Center 292
 - international treaties 287
 - International Union of Geodesy and Geophysics (IUGG) 36, 207, 285, 286, 289, 290
 - International Union of Radio Science (URSI) 16, 202, 207, 227
 - International VLBI Service (IVS) 290
 - International VLBI Service for Geodesy and Astrometry (IVS) 285
 - International VLBI Service for Geodesy and Astronomy (IVS) 266, 269, 289, 293
 - Internet 255
 - invariant interval 98
 - ion 153
 - ion trap 176
 - Paul 176, 183
 - Penning 177
 - ionosphere 251, 268
 - ionospheric delay 271
 - ionospheric map 271
 - ionospheric variations 256
 - IPMS, *see* International Polar Motion Service
 - ITRF, *see* International Terrestrial Reference Frame
 - ITRS, *see* International Terrestrial Reference System
 - ITU, *see* International Telecommunications Union
 - IUGG, *see* International Union of Geodesy and Geophysics
 - IVS, *see* International VLBI Service or International VLBI Service for Geodesy and Astrometry
- j**
- J2000.0 118
 - Jason-1 276, 295
 - Jeffreys, Harold 52
 - Jet Propulsion Laboratory (JPL) 29, 43, 122, 176
 - John of Saxony 25
 - JPL, *see* Jet Propulsion Laboratory
 - Jupiter 110, 238, 239
 - satellites 46
- k**
- Kant, Immanuel 41, 52
 - Kelvin, Lord 52
 - Kepler's laws 5, 26, 237
 - Kepler, Johannes 26, 27
 - Knibb, Joseph 136
 - Königsberg, Johannes Müller von 25
 - Küstner, Karl Friedrich 67
 - Kusch, P. 158
- l**
- Laboratoires Suisse de Recherches Horlogères 201
 - LAGEOS 273, 276
 - LAGEOS-1 294
 - LAGEOS-2 294
 - Lalande, Jérôme 28, 42
 - Landsbergen 26
 - Laplace 30, 42
 - laser 272, 299
 - laser ranging 75, 242, 264, 293, 309
 - LCT, *see* Local Civil Time
 - leap second 228, 229, 232, 245, 307, 308
 - LED, *see* light-emitting diode
 - Leibnitz, Gottfried Wilhelm 32
 - length 240
 - contraction 99
 - proper 253
 - length of day (LOD) 51, 54, 88, 89, 232, 270, 272, 276, 278
 - Lense-Thirring effect 108
 - LeVerrier, Urbain Jean Joseph 28
 - Lévy, M. Jacques 80
 - librational motions 31
 - light deflection 34
 - light-emitting diode (LED) 299
 - limb 31
 - limb corrections 86
 - Linear Ion Trap Frequency Standard (LITS) 176
 - Liouville equation 69
 - Lissajous, Jules 139
 - LITS, *see* Linear Ion Trap Frequency Standard
 - LLR, *see* Lunar Laser Ranging
 - Local Civil Time (LCT) 10
 - local frame 103
 - LOD, *see* length of day
 - Longitude Act of Great Britain 138

LORAN 227, 243, 298
 LORAN-C 201, 227, 257
 Lorentz transformation 98, 101
 Lunar Laser Ranging (LLR) 34, 88, 242, 294
 lunar orbit 52
 lunar theory 5, 29, 30, 43, 85
 lunar tides 53
 Lyons, Harold 157

m

magnesium 177
 magnetic field 53
 Manfredi, Eustachio 27
 Markowitz Moon Camera 193, 195
 Markowitz, William 85, 87, 162, 192
 Mars 45, 47, 49
 – time 247
 mass 240
 – gravitational 104, 106
 – inertial 104
 Mayer, Johann Tobias 28, 30, 42
 Mayer, Julius Robert 52
 MCXO, *see* oscillator, micro-computer-controlled crystal
 Mercury 44, 45, 49, 50, 53, 90, 238
 – meridian observations 47
 – perihelion 37, 46
 – transits 46, 47
 mercury 175, 176, 183, 184
 Merit Working Group 76
 meter 196, 197, 241, 308
 – definition 241
 – SI 118
 Meter Convention 281, 283
 metre 240
 metric 37, 103, 104, 120, 121, 219
 metric potentials 108
 metric tensor 105, 106, 109
 metrology 282
 Michelson and Morley 96
 microwave 152
 microwave energy 156
 microwave frequencies 181
 microwave links 254
 Minkowski diagram 99–102
 Minkowski metric 98
 minute 136, 189–191
 models
 – ionospheric 259
 – tropospheric 259
 modified Allan variance 148
 modified Hadamard variance 148

moment of inertia 53
 Montebruni, Francesco 27
 month 2
 – synodic 30
 Moon 18, 23, 24, 27–30, 32, 34, 36, 37, 41–43, 46, 53, 64, 69, 83, 85, 86, 89, 90, 192, 237–242, 247, 274
 – crescent 239
 – ephemeris 85
 – librations 34
 – longitude 45, 48–50
 – mean anomaly 73
 – mean longitude 31, 42, 73, 88
 – mean motion 44
 – motion 30, 42
 – orbit 34, 53
 – orbital motion 53
 – secular acceleration 85
 moonrise 239
 moonset 239
 Moscow Time 258
 motions
 – luni-solar 73
 – planetary 73
 Mu'ayyad al-Din al-'Urdi 25

n

NASA 273
 Nasis al-Din Tusi 25
 National Bureau of Standards (NBS) 200, 201, 226, 227
 National Institute of Standards and Technology 226
 National Physics Laboratory (NPL) 158, 193, 199, 201, 226
 National Research Council of Canada 200
 National Research Laboratory of Canada (NRC) 201
 Nautical Almanac 28, 223, 225
 Nautical Almanac Offices 32
 – British 28
 – U.S. 28
 Naval Research Laboratory 200
 navigation 298, 305, 308
 navigation satellite 258
 navigation signals
 – low-frequency 257
 navigation system 242, 243
 – hyperbolic 243
 navigational satellite systems 271
 navigational satellites 264
 navigational system 245
 NBS, *see* National Bureau of Standards

- NBS-A 200
- Neptune 28, 33
- Network Time Protocol (NTP) 255, 300, 301
- neutron star 247
- Newcomb's Tables
 - of the Sun 31, 35, 83
 - of the Sun, Mercury and Venus 50
- Newcomb's Theory of the Sun 13, 79, 86
- Newcomb, Simon 19, 31, 35, 36, 43–46, 71, 90, 91
- Newton's law of gravitation 26, 237
- Newton's Principia 27
- Newton's universal law of gravity 26, 237
- Newton, Isaac 2, 26, 27, 30, 41, 79
- Newtonian gravitational theory 95
- Newtonian mechanics 103, 113
- Newtonian potential 103, 105
- Newtonian space time 103
- Nicholson, Alexander McLean 140
- noise
 - clock 146
 - flicker frequency 146, 147
 - flicker phase 146–148
 - random walk frequency 146, 147
 - spectral density 147
 - thermal 170
 - white frequency 146, 147
 - white phase 146–148
- nonrotating origin 15
- NPL, *see* National Physics Laboratory
- NRC, *see* National Research Laboratory of Canada
- NRL, *see* Naval Research Laboratory
- NTP, *see* Network Time Protocol
- nucleus 153, 155
- numerical integration 33, 91
- nutation 12, 13, 15, 31, 61–66, 69, 70, 72, 73, 76, 92, 118, 263
- Nyrén, M. 67
- o**
- OAM, *see* Oceanic Angular Momentum
- observational techniques 237
- observations 57
 - ancient eclipse 48
 - laser 35
 - lunar 35
 - lunar laser ranging 87
 - radar 35
 - spacecraft 34, 35
 - transit 45
 - VLBI 87
- Observatoire de Neuchâtel (ON) 200, 201
- Observatoire de Paris 80, 201, 291, 292
- occultation 32, 33, 50, 86, 88, 237–239
- ocean and atmospheric loading 272, 273
- Oceanic Angular Momentum (OAM) 54, 278
- OCXO, *see* oscillator, oven-controlled crystal
- Omega 227
- ON, *see* Observatoire de Neuchâtel
- optical atomic standards 181
- optical fiber 251, 254, 299, 309
- optical frequencies 157
- optical frequency standard 184
- optical lattice 184, 185
- optical pumping 174, 175
- optical standards, characterizing 186
- optical transition frequencies 181
- orbit 26
- orbital elements 27
- orientation 57
- origin of longitudes 62, 64
- oscillator
 - micro-computer-controlled crystal (MCXO) 151
 - oven-controlled crystal (OCXO) 151
 - quartz-crystal 151, 154, 157
 - temperature-compensated crystal (TCXO) 151
- p**
- parallax 5, 12, 31
- Parameterized Post-Newtonian (PPN) 37
 - formulations 92
- Paris Observatory 16, 27, 207, 226, 227
- Parry, J. V. L. 158, 192
- Pascal, Blaise 32
- perihelion 18
- Peters, C.A.F. 67
- phase measurements 257
- Philolaus 41
- Photographic Zenith Tubes (PZT) 56, 75, 76, 84, 193, 194
- Physikalische Technische Bundesanstalt (PTB) 201
- piezo-oscillator 192
- piezoelectric activity 139
- piezoelectricity 140
- Planck's constant 152
- planet 24, 27, 41
- planetary masses 36
- planetary theories 79
- planets 26, 34, 36, 37, 48, 49, 235

polar motion 61–67, 69–72, 75, 76,
270–272, 274, 276, 278, 279, 294
pole, free motion 66
Pontecoulant, Gustave de 30
Pope Gregory XIII 2
positions
– mean 62
– terrestrial 62
– true 62
post-glacial rebound 53, 89
post-Newtonian parameters 34, 308
post-Newtonian potential 108
Post-Newtonian Potential Coefficients 109
potential 124, 253
potential difference 252
power grid 300
PPN, *see* Parameterized Post-Newtonian
precession 12, 13, 15, 31, 61–66, 69, 70,
72, 76, 83, 90, 118, 263
– general 73
– planetary 73
precession-nutation 64, 66, 70, 112, 269
precise positioning 305
precision 143, 302, 304
predictions 279
primary frequency standard 206, 215,
216, 220
primary standards 215
prime meridian 224
propagation 255
propagation delay 250
propagation effects 250
proper motions 15, 35
PTB, *see* Physikalische Technische
Bundesanstalt
Ptolemy 3, 9, 24, 25, 35, 189, 236
pulsar 247, 285, 307
– millisecond 248, 307
– stability 248
– timing 247, 248
punched card equipment 32
PZT, *see* Photographic Zenith Tubes

q

quadrupole moment 252
quality (Q) factor 142
quantum entanglement 309
quantum logic standard 185
quantum number 153
quartz 140, 141, 301
quartz crystal 158, 200
quartz oscillator 152, 171, 172, 299, 300
quasar 265, 266

r

Rabi Pedestal 164
Rabi, Issac 158
radar observations 34
radar ranging 241
radio regulations 287, 288
radio signals
– high-frequency 255
– low-frequency 256
radio source position 265
radio sources 269
radio station, very-low-frequency 200
radio telescope 265
Radiocommunications Conferences 288
Ramsey Fringe 164
Ramsey, Norman 158
reference frame 3, 36, 57, 61, 99, 102,
117, 285
– astronomical 34
– barycentric 37
– celestial 118, 265
– celestial barycentric 118
– dynamical 34
– geocentric 119, 204
– geocentric nonrotating 253
– geocentric rotating 37
– inertial 95, 96, 104
– nonrotating 37
– nonrotating local inertial 252, 253
– rotating 252, 253
– terrestrial 272, 274
reference system 90
– barycentric 109, 111, 121
– barycentric celestial 105
– celestial 10, 19, 35, 61, 72, 117,
263
– celestial barycentric 118
– dynamical 34
– geocentric 109, 111
– geocentric celestial 105
– Newtonian 95
– star catalog 32
– terrestrial 19, 61, 72, 263
refraction 5, 12
Reinhold, Erasmus 25
relativistic corrections 260
relativistic effects 33, 34, 37, 208, 245,
248, 250, 251, 272, 273
relativistic frequency shift 215
relativistic theories 36
relativistic time comparison 252
relativistic time scale 247
relativity 2, 95, 122, 236, 308

- general 95, 98, 103–105, 115, 118, 119, 205, 294
- special 95, 98, 99, 103, 104
- requirements 297
- resonance frequency 156, 172, 174
- resonance phenomenon 142
- retro reflector 242, 272, 273
- RGO, *see* Royal Greenwich Observatory
- Riefler escapement 137
- Riemannian space time 103
- right ascensions 3, 35, 64, 69
- Ross, Frank 45
- Royal Greenwich Observatory (RGO) 30, 160, 192, 199, 201, 226
- Royal Observatory 137
- rubidium 153, 172, 173, 301
- rubidium cells 172, 173
- rubidium fountains 173
- rubidium maser, double-bulb 173
- rubidium standard 299, 300
- Rudolphine Tables 26

S

- Sagnac correction 246
- Sagnac delay 245
- sand glass 190, 191
- Satellite Laser Ranging (SLR) 272, 274
- satellite navigation systems 257
- Saturn 110, 238
- Scheutz, Edvard 32
- Scheutz, Georg 32
- scientific unions 283
- seasonal variation 71
- second 9, 120, 130, 136, 161, 189–191, 195, 197, 202, 215, 240, 241
 - atomic 87, 228
 - definition 162, 191, 308
 - ephemeris 84, 86, 87, 91, 115, 162, 191, 192, 195
 - ephemeris time 90
 - historical 189
 - mean solar 82, 87
 - SI 87, 98, 114, 115, 117, 118, 120, 123, 124, 158, 192, 193, 195, 196, 205, 207, 215, 216, 218, 220, 228, 231, 232, 241, 308
 - solar 90
 - TAI 228
 - universal time 87
- secular acceleration 7, 31, 49, 229
- secular polar motion 70
- Seleukus 41
- semi-diurnal variation 245

- sexagesimal system 3, 130
- Shortt clock 138
- Shortt, William H. 138
- SI units 197, 282, 283
- SI, *see* International System of Units or *Système international d'unités*
- signal 258
- simultaneity 100
- single ion 184
- Sirius 130
- SLR, *see* Satellite Laser Ranging
- solar corona 34
- solar flares 251
- solar pressure 272, 273
- solar system 23, 24, 29, 61, 65, 79, 90, 91, 104, 105, 107–109, 113–115, 119, 121, 124, 235–237, 247, 248
- solar system phenomena 237
- solstices 18
- SONET 299
- sources, extended structure 269
- space-time coordinates 118
- special relativity 95, 98, 99, 103, 104
- spectrum management 299
- speed of light 96, 213, 239, 241, 251, 252, 269
- Spencer Jones, Harold 47, 50, 53, 81, 82
- SPOT-2 276, 295
- SPOT-3 276, 295
- SPOT-4 276, 295
- SPOT-5 276, 295
- spread spectrum 259
 - technique 299
- stability 144, 177, 186, 205, 214, 217
- star catalog 5, 35, 85, 86, 91, 238
 - FK4 35
 - FK5 35
 - fundamental 35
- Starlette 273
- Stella 273
- stepped atomic time 227
- Sterne, T.E. 53
- stored-ion devices 176
- Stoyko, Nicholas 51
- stratum level 299, 300
- Streete, Thomas 27
- strontium 184
- Struve, F.G.W. 67
- Struve, O.W. 67
- Struve, Wilhelm 3
- Study Group 7 288

Sun 23, 24, 26–28, 33, 36, 37, 41–43,
45–50, 53, 57, 64, 68, 69, 81, 83, 85, 86,
90, 110, 129, 130, 161, 237, 238, 240
– mean anomaly 73
– mean longitude 124, 191
Sun and planets 29
sundial 5, 6, 129, 239
sunrise 239
sunset 239
sunspot numbers 45
synchronization 142
syntonization 142
Sysiphus cooling 167
system, heliocentric 25
Système international d'unités (SI) 202,
281, 282

t

TA(BIH) 201
TAI, *see* International Atomic Time
TAI-UTC 230
TA(k) 201, 208
TCA, *see* Areocentric Coordinate Time
TCB, *see* Barycentric Coordinate Time
TCG, *see* Geocentric Coordinate Time
TCXO, *see* oscillator, temperature-
compensated crystal
TDB, *see* Barycentric Dynamical Time
TDM, *see* Time Division Multiple Access or
Time Division Multiplexing
TDT, *see* Terrestrial Dynamical Time
TEC, *see* Total Electron Content
tectonic plate motion 269, 274
tectonics 285
telegraph 249
telephone 254
telescopic observation 264
television broadcast 255
Terrestrial Dynamical Time (TDT) 92, 114,
115, 124, 219
Terrestrial Intermediate Origin (TIO) 15,
64, 68
Terrestrial Intermediate Reference System
64
Terrestrial Reference System (TRS) 64,
264
Terrestrial Time (TT) 18, 54, 110, 111,
119–126, 216, 218–221, 247, 251
theory
– lunar 88
– of relativity 29, 37, 79, 91, 92, 103
Third World Academy of Sciences (TWAS)
289

tidal acceleration 89
tidal bulge 42
tidal deceleration 42, 44, 49, 65, 90
tidal dissipation 89
tidal friction 5, 30, 43, 49, 52, 53, 90
tidal interactions 31
tidal potential 108
tidal secular acceleration 88
tidal variations 194
tide-generating potential 253
tides 7, 54, 269
– Earth 272, 273
– ocean 52, 67, 72, 272, 273
– solid Earth 52
time 104, 240, 251
– absolute 2
– accuracy 236
– apparent sidereal 12
– apparent solar 7, 10, 11
– local 7
– applications 297
– astronomical 55, 56, 76, 79
– atomic 36, 37, 57, 86, 87, 90, 91, 113–115,
117, 124, 162, 199, 201, 202, 207, 247, 248,
307
– balls 6, 249
– barycentric 114
– civil 228
– comparison 208
– concepts 1
– coordinate 91, 98, 99, 102, 109,
112, 113, 115, 118–120, 122–126, 204,
208, 213, 251, 253
– dilation 97, 99, 101, 102, 104, 156,
245
– dissemination 250, 254, 281
– distribution 232, 308
– dynamical 36, 37, 80, 82, 92, 111,
113–115, 124, 126
– ephemeris 82, 121, 162
– epoch 5
– future 307, 308
– geocentric 114
– Greenwich mean sidereal 14
– Greenwich mean solar 231
– local 6
– mean solar 10, 11, 13, 15, 34, 86, 191,
228, 239
– Newtonian 46, 47, 55, 80, 81
– proper 91, 98, 99, 102, 115, 118–120,
123, 246, 252
– pulsar based 248
– relativistic 122

- rotation 47
- scale 1, 23, 36, 37, 56, 61, 79, 80, 84, 87, 91, 92, 111, 113–115, 117–119, 122, 124–126, 160, 199–202, 204, 206, 207, 236, 247, 248, 281, 307, 308
- sidereal 12
- signal 6, 91, 194, 226, 249, 250, 256, 257
- solar 7, 9
- space mission 246
- stamp 301
- transfer 111, 201, 208, 249, 251, 258
- transformations 109
- uniform 2, 35, 55–57, 79, 113, 160, 236, 237
- unit 1, 190
- variation 148
- zone 20, 224, 232
- time and frequency needs 301
- Time Division Multiple Access (TDM) 299
- Time Division Multiplexing (TDM) 298
- Timation 243
- time standards, atomic 219
- TIO, *see* Terrestrial Intermediate Origin
- Tompion, Thomas 137
- TOPEX/POSEIDON 276, 295
- Total Electron Content (TEC) 268
- total Hadamard variance 148
- total variance 148
- transit 237, 238
- transit circle observations 33
- transit circles 56, 75, 76, 84–86
- transition frequency 155, 175
- transportation 305
- trap
 - ion 174, 175
 - Paul 174, 175, 185
 - Penning 174
 - shuttle 176
- Treaty of the Meter 281, 283
- TRF 75
- tropical year 191
- tropospheric and ionospheric conditions 255
- troposphere 34
- tropospheric delay 272
- TRS, *see* Terrestrial Reference System
- TT, *see* Terrestrial Time
- TWAS, *see* Third World Academy of Sciences
- two-way satellite time and frequency transfer (TWSTFT) 208, 259, 260

U

- UNESCO, *see* United Nations Educational, Scientific and Cultural Organization
- uniformity 236
- Union Radio-Scientifique Internationale (URSI) 289
- United Nations Educational, Scientific and Cultural Organization (UNESCO) 289
- universal day 225
- universal gravitation 5
- Universal Time (UT) 10, 14, 15, 17, 47, 82, 84, 91, 192, 193, 223, 226
- universe 235
- University of Luxembourg 292
- University of Texas 242, 243, 272
- Uranus 28, 33
- URSI, *see* International Union of Radio Science or Union Radio-Scientifique Internationale
- U.S. Naval Observatory (USNO) 6, 125, 162, 193, 199, 201, 226, 245, 291, 292
- USNO, *see* U. S. Naval Observatory
- UT, *see* Universal Time
- UT0 14, 17, 226
- UT1 14–17, 54, 68, 76, 192, 226, 228, 231, 245, 308
- UT1-UTC 65, 68, 75, 76, 264, 269, 270
- UT2 14, 16, 17, 160, 193, 195, 200, 201, 226–228
- UT2 second 194
- UTC, *see* Coordinated Universal Time
- UTC(k) 218

V

- variations
 - fortnightly 51
 - monthly 51
- Venerable Bede 189
- Venus 45, 47, 49, 50, 53, 90, 238, 239
- verge escapement 132, 135
- Very Long Baseline Interferometry (VLBI) 15, 16, 68, 76, 264–270, 293
 - observation 271
- visual zenith telescopes 76
- Viviani, Vincenzo 133
- VLBI, *see* Very Long Baseline Interferometry
- VLF 201

W

- water utilities 301
- Watts, C.B. 86
- Williams, N.H. 157
- Winkler, Gernot 16, 228

- Working Party 7A 288
- World Administrative Radio Congress 226
- World Data Centers 290
- World Data System 290
- world line 100
- World Radiocommunications Conferences (WRC) 288
- WRC, *see* World Radiocommunications Conferences
- WWV 193, 226
- Y**
 - year 2
 - Julian 19
 - tropical 18, 19, 83
 - ytterbium 177, 184
- Z**
 - Zacharias, J.R. 158
 - Zeeman effect 153
 - zenith telescopes 75
 - zijes 25