Astrophysics and Space Science Proceedings 50

Elisa Felicitas Arias
Ludwig Combrinck
Pavel Gabor
Catherine Hohenkerk
P. Kenneth Seidelmann *Editors*

The Science of Time 2016

Time in Astronomy & Society, Past, Present and Future

Springer

Astrophysics and Space Science Proceedings

Volume 50

More information about this series at http://www.springer.com/series/7395

Elisa Felicitas Arias • Ludwig Combrinck
Pavel Gabor • Catherine Hohenkerk
P. Kenneth Seidelmann
Editors

The Science of Time 2016

Time in Astronomy & Society, Past, Present and Future

Springer

Editors
Elisa Felicitas Arias
Time Department
International Bureau for Weights and Measures
Sevres, France

Pavel Gabor
Vatican Observatory
Tucson, AZ, USA

P. Kenneth Seidelmann
Department of Astronomy
University of Virginia
Charlottesville, VA, USA

Ludwig Combrinck
Hartebeesthoek Radio Astronomy Observatory
Krugersdorp, South Africa

Catherine Hohenkerk
HM Nautical Almanac Office
UK Hydrographic Office HM Nautical Almanac Office
Taunton, Somerset, UK

ISSN 1570-6591 ISSN 1570-6605 (electronic)
Astrophysics and Space Science Proceedings
ISBN 978-3-319-59908-3 ISBN 978-3-319-59909-0 (eBook)
DOI 10.1007/978-3-319-59909-0

Library of Congress Control Number: 2017946180

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The symposium "The Science of Time" was held on June 5–9, 2016, at Harvard University in Cambridge, Massachusetts, USA. The symposium was to consider all scientific aspects of time.

The uses of time in astronomy—from steering and pointing telescopes, coordinating and processing observations, predicting ephemerides, determining Earth orientation, and analyzing time series data to many other ways—represent a broad sample of how time is used throughout human society and in space. Time and its reciprocal, frequency, are the most accurately measurable quantities and often important paths to the frontiers of science. But the future of timekeeping is changing with the development of optical frequency standards and the resulting challenges of distributing time at ever-higher precision, with the possibility of timescales based on pulsars, and with the inclusion of higher-order relativistic effects. The definition of the second will likely be changed before the end of this decade, and its realization will increase in accuracy; the definition of the day is no longer obvious. The variability of the Earth's rotation presents challenges of understanding and prediction. It is time to take a closer look at time in astronomy and other sciences as a defining element of modern civilization.

The symposium aimed to set the stage for future timekeeping standards, infrastructure, and engineering best practices for astronomers and the broader society. At the same time, the program was cognizant of the rich history from Harrison's chronometer to today's atomic clocks and pulsar observations. The theoreticians and engineers of time came together with the educators and historians of science, enriching the understanding of time among both experts and the public.

"The Science of Time" was hosted by the Harvard-Smithsonian Center for Astrophysics (CfA) in Cambridge, Massachusetts, USA; the CfA is a collaboration of the Smithsonian Astrophysical Observatory (SAO) and the Harvard College Observatory (HCO). Some CfA staff members are in the Harvard Department of Astronomy (HDA).

The following symposium topics are included in these proceedings:

- The scientific and technical uses of time and time series data
- The civil and scientific understanding of time—education and outreach
- The history of time and timepieces, clocks, and calendars
- Social, cultural, and religious uses of timing information
- Time from sundials and the pendulum to atomic clocks and pulsars
- Impact of precise time and frequency measurement in astronomy and science
- Earth rotation and time
- Time and solar-system ephemerides
- The physics of time

The proceedings contain many papers from the symposium. Where papers were already published or not available, abstracts and references to published papers, where available, are included.

Sevres, France — Elisa Felicitas Arias
Krugersdorp, South Africa — Ludwig Combrinck
Tucson, AZ, USA — Pavel Gabor
Taunton, Somerset, UK — Catherine Hohenkerk
Charlottesville, VA, USA — P. Kenneth Seidelmann

Contents

Chapter 1
Cosmic Time: From the Big Bang to the Eternal Future

Chris Impey

Abstract Cosmology presents intriguing issues for the understanding and tracking of time. The big bang theory says that the universe began 13.8 billion years ago, in a situation of almost unimaginable temperature and density. The age of the universe is highly constrained by the world model, as long as there are reliable measurements of the current expansion rate (the Hubble constant), and amounts of baryonic matter, dark matter, and dark energy. A crucial cross-check on the model age comes from stellar chronometers and the ages of the oldest stars in the Milky Way. Landmarks in cosmic evolution reach back to about 380,000 years (recombination), 40,000 years (matter domination), and a few minutes (light element creation) after the big bang.

Time in Newtonian cosmology is absolute, linear, and eternal, whereas time in the modern cosmology is governed by Einstein's general relativity, a geometric theory which embodies a profound connection between space and time. In relativity, objects travel on paths called world lines in four-dimensional spacetime. Relativity defines proper time as time measured by an observer with a clock following a world line. A clock in motion relative to the observer, or in a different gravity situation, will not measure proper time. The time concept rests on the cosmological principle—the assumption that the universe is homogeneous and isotropic on large scales. If that is true, there are well-defined relationships between time, scale factor, and temperature going all the way back to the first fraction of a second after the big bang. Time in the far future of the universe can be measured in terms of physical processes—the spinning down of pulsars and the evaporation of black holes. There are still profound physical and philosophical issues raised by the definition of clocks and observers in cosmology.

Keywords Cosmology • Universe • Cosmic time • Big bang • Expansion

C. Impey (✉)
Department of Astronomy, University of Arizona, Tucson, AZ 85721, USA
e-mail: cimpey@as.arizona.edu

E.F. Arias et al. (eds.), *The Science of Time 2016,* Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_1

Time and Entropy

This paper will present a brief overview of time in modern cosmology, the subject that deals with the contents, the geometry, the history, and the future of the universe. Some complex topics will inevitably be omitted or covered lightly. Books have been written on the philosophy of time, and regrettably that fascinating subject will also not be covered in this paper (Mauldin 2013; Smeenk 2013). The underpinning of cosmology is the fact that the universe evolves; it began in a situation of extremely high density and temperature, and it is inexorably expanding toward an enormous state of extreme cold and almost-perfect vacuum.

Physical time embodies an interesting conundrum. Microscopic processes operate equally well forward and backward in time. Except for a very slight time asymmetry that manifests in the joint operation of charge and parity, they're invariant under any time reversal. This applies to the operations on time and linear momentum, $t \rightarrow -t$ and $p \rightarrow -p$, in Newton's gravity:

$$\frac{dp}{dt} = -\frac{GMm}{r^2}$$

However, the second law of thermodynamics says that entropy tends to increase in any closed system and is not invariant under time reversal:

$$\frac{dS}{dt} \geq 0, \quad S = k \log W$$

where S is the entropy of the system, k is Boltzmann's constant, and W is the number of equivalent microscopic states. In systems of many particles, there are many more disordered or high-entropy states than ordered or low-entropy states, which causes entropy to increase as microscopic interactions occur in equilibrium. The familiar example is the irreversible mixing of two gases that had previously been separated. There have been some clever (but not entirely successful) attempts to argue that time asymmetry arises at the microscopic level (e.g., Maccone 2009).

At first glance, the evolution of the universe seems to violate the second law. As it evolves, the structure goes from the simplicity and near uniformity of a hot gas just after the big bang to a set of stars condensed by gravity and aggregated into galaxies and clusters of galaxies. However, as the universe evolves, interaction rates reduce faster than the expansion rate, so processes go out of equilibrium. If the gravitational degrees of freedom are tracked, the total entropy increases (Fig. 1.1).

A nonastronomical example of small-scale organization and structure within the larger context of increasing entropy is biology. Living creatures organize materials for life functions and emit heat, thereby raising the entropy of their surroundings. But life is only possible because it exists at a temperature of 300 K and mediates a transaction between the much hotter Sun at 5700 K and the heat sink of deep and mostly empty space at 2.7 K (Avery 2003; Balbi 2013).

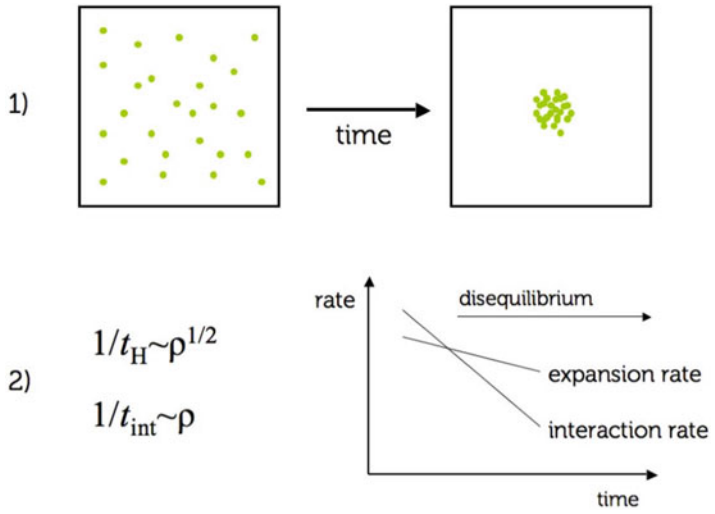


Fig. 1.1 As the universe evolves over time, structures form, in apparent violation of the second law of thermodynamics. But processes go out of equilibrium due to a rapid decrease of the interaction rate, so when the number of gravitational degrees of freedom is properly accounted for, the entropy is increasing (courtesy of Balbi 2013)

Time in Relativity

In Newtonian gravity, space and time are linear, absolute, and infinite. Timekeeping does not depend on location or motion, and synchronized clocks stay synchronized forever (Belot 2013). The universe evolves at the same rate at all locations.

Things get more complicated with relativity. The finite speed of light leads to the concept of a light cone, a surface that defines the limit of causal communication in the usual two-dimensional representation of space with an orthogonal time axis. An observer cannot be affected by events outside their past light cone, and they cannot influence events outside of their future light cone (Fig. 1.2). In the big bang model, the simple light cone for static space does not apply. Spacetime has expanded at a continuously varying rate since the big bang and the expansion was superluminal early in cosmic history (and for all points outside the Hubble sphere, defined by the radius c/H_0, where H_0 is the Hubble constant or current expansion rate). This means that in a 13.8-billion-year-old universe, regions of space as far away as 46 billion light years in any direction are within our horizon (Davis and Lineweaver 2004).

Within the framework of special relativity, Hermann Minkowski introduced the idea of proper time, denoted τ. Proper time is the elapsed time between two events as measured by a clock passing through both events (Cook 2004). It depends on the events and also on the motion of the clock between events. By contrast, coordinate time, t, is the apparent time between two events as measured by a distant observer using their own clock (Fig. 1.3).

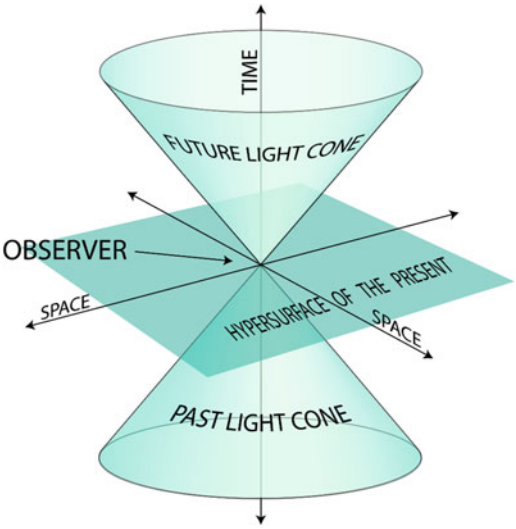


Fig. 1.2 With the third dimension suppressed, a light cone shows the surface of all possible light signals arriving at and departing from a point in static spacetime. The finite speed of light sets a context for causality and signal propagation in the universe (Stib, Wikipedia Creative Commons)

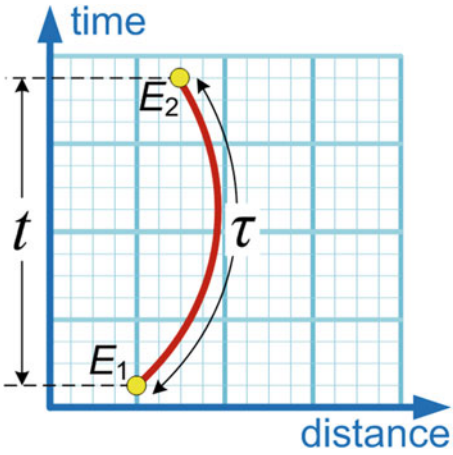


Fig. 1.3 The vertical axis is the coordinate time, t, of an inertial observer measuring an interval between the events E_1 and E_2. The proper time, τ, is time measured by a clock passing through both events and so is independent of coordinates and a Lorentz scalar (Dr. Greg, Wikipedia Creative Commons)

We can see the important role of clocks in the measurement of time. But even if we imagine atomic clocks that function reliably, independent of time and space, there is a special measurement problem with the universe as a whole, since there was a time when there were no atoms and the concept of a distant observer is meaningless. As Misner et al. (1973) framed it in their magisterial tome *Gravitation*, "Each actual clock has its ticks discounted by a suitable factor—3×10^7 s per orbit for the Earth-Sun system, 1.1×10^{-10} s per oscillation for the Cesium atom, etc. Since no single clock (or finite size and strength) is conceivable back to the singularity, statements about proper time since the singularity rely on an infinite sequence of smaller and sturdier clocks with the ticks discounted and added."

Special relativity defines inertial reference frames for uniform relative motion. An inertial frame describes time and space homogenously (the same at any location) and isotropically (the same in any direction). Clocks moving with respect to any inertial frame run more slowly (Taylor and Wheeler 1992), and time dilation is described by the Lorentz transformation, where T_0 is time measured in the inertial frame and T is time measured by someone moving at velocity v:

$$T = T_0 / \sqrt{1 - v^2/c^2}$$

With general relativity, Einstein extended the framework to accelerated motion and applied it to gravity, where accelerated motion due to gravity cannot be distinguished from accelerated motion due to any other force (which is the equivalence principle). A gravitational field causes time dilation; in a stronger gravity field, a clock will run slower. In special relativity the effect is reciprocal, while in general relativity, two observers will agree on the size and direction of the effect, measured at a distance R from mass M, where G is the gravitational constant, given by:

$$T = T_0 / \sqrt{1 - 2GM/Rc^2}$$

Time dilation was first detected by comparing two atomic clocks, one of which was flown in a high-altitude jet and the other of which remained on the ground. It has recently been measured with exquisite accuracy using optical atomic clocks (Chou et al. 2010). In this formalism, a photon loses energy moving in a gravity field, which is the gravitational redshift. For a compact object, the denominator diverges when $R = 2GM/c^2$, defining the event horizon of a black hole, where time dilation becomes infinite.

Relativistic Cosmology

Hubble showed that the recession velocities of galaxies could be explained in terms of an expanding universe and the linear expansion sets an approximate age for the universe, $v = H\,d$ and $t = 1/H$, where v is the recession velocity for a galaxy at a distance d and H is the Hubble parameter, or expansion rate. In the local universe $H_0 = 70$ kms^{-1} Mpc^{-1}, which gives $t = 14$ Gyr (14 billion years) by simply projecting a linear expansion back to zero separation.

Accurate calculation of the age of the universe requires general relativity, where the geometry and evolution of spacetime are governed by the mass-energy distribution. In general, it is difficult to define a universal coordinate system (x, y, z, t). But if gravity is cancelled using free-falling observers, also called fundamental observers, it is possible to use special relativity to define a local coordinate system (relying on the equivalence principle). A free-falling observer measures the spacetime of special relativity and its line element as defined by Minkowski:

$$-ds^2 = c^2 d\tau^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

All fundamental observers will measure time changing at the same rate, *dt*, and there is a universal time coordinate, *t*. Synchronizing universal time everywhere requires an assumption or symmetry principle. In cosmology (as in special relativity above), the assumption is isotropy, which is bolstered by measuring the same Hubble flow, large-scale galaxy distribution, and microwave background radiation temperature in different directions, and homogeneity, which cannot formally be tested since data at different distances are comparing different epochs. The assumption is a statement of the Copernican principle that we're not in a special location in the universe. A few studies have suggested the presence of large and potentially anomalous structures, but the largest surveys demonstrate homogeneity (Scrimgeour et al. 2012; Clowes 2013; Horvath et al. 2014).

The assumption of homogeneity and isotropy is called the cosmological principle, and it is foundational in big bang cosmology. Under this assumption, the density is the same everywhere at any time, and the density and radiation temperature decrease smoothly with time after the big bang. In theory, all fundamental observers can have their clocks synchronized at a particular time and density, $\rho\,(t_0) = \rho_0$. In relativistic cosmology, with spacetime described by general relativity and assuming symmetry, the Minkowski line element can be generalized into the Robertson-Walker metric (Ibison 2007):

$$-c^2 d\tau^2 = -c^2 dt^2 + a(t)^2 \left(\frac{dr^2}{1 - kr^2} + r^2 d\phi^2 \right)$$

where a is the scale factor, or the distance between any two points in space as a function of time, r and ϕ describe the polar coordinate system, and k is the curvature of space, with $k = 0$ for flat and infinite space, $k = 1$ for closed and finite space, and $k = -1$ for open and infinite space (Peacock 1998). The fundamental observable of cosmology is the redshift, z, defined by $(1 + z) = a_o/a_e = \lambda_o/\lambda_e$, the ratio of the scale factor when a photon is observed to the scale factor when it was emitted.

Cosmic Evolution

The evolution of the universe is governed by the mass and energy it contains. As the universe expands, the time evolution of the size and its dependence on density for a spatially flat model is:

$$R \propto t^{\frac{2}{3}}, \qquad t = 1/\sqrt{6\pi G \rho_m}$$

where R is the scale factor or distance between any two points and ρ_m is the matter density. Since $\rho_m = \rho_0\,(R/R_0)^{-3}$, radiation density, ρ_r, declines at a faster rate, $\rho_r = \rho_0\,(R/R_0)^{-4}$, since it depends on volume with an extra stretching factor of

redshift. Before 40,000 years after the big bang, the universe is radiation dominated, and the time evolution of the size and its dependence on density is:

$$R \propto t^{\frac{1}{2}}, \qquad t = 1/\sqrt{32\pi G\rho_r}$$

Matter caused the cosmic expansion to decelerate for the first half of the time since the big bang. Matter-vacuum equality occurs at ¾ of the current age, while transition from deceleration to acceleration occurs earlier, at 56% of the current age. In the cosmic "pie chart" of constituents, about a third is matter, with 5% as baryons and 27% as dark matter (Spergel 2015). However, in the last 5 billion years, dark energy has been causing an accelerating expansion. Dark energy constitutes 68% of the universe (Fig. 1.4). Due to its peculiar property of negative pressure, dark energy causes an exponential increase in the scale factor, $R = R_0\, e^{Ht}$.

The age of the universe in the big bang model comes from integrating the expansion over cosmic history (Liddle 2003). The integration can be done numerically with an approximation:

$$t_0 = \int_0^\infty \frac{dz}{(1+z)H(z)} \cong 2/3H_0(0.7\Omega_M + 0.3(1-\Omega_\Lambda))^{-0.3}$$

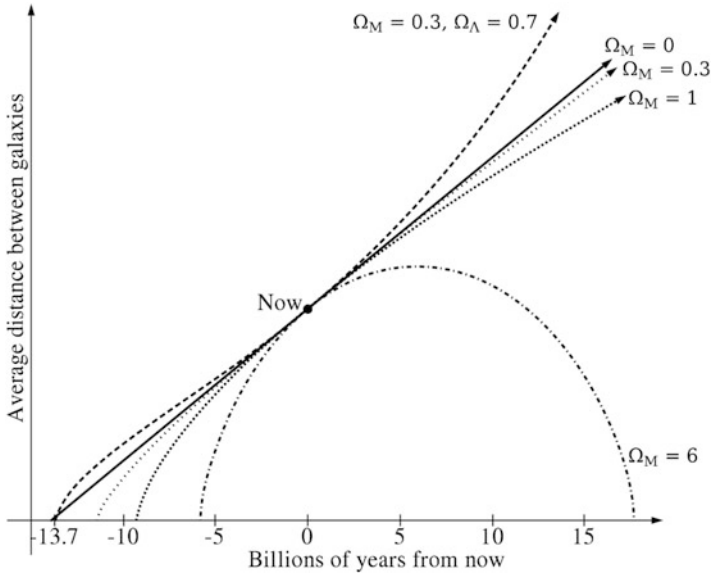


Fig. 1.4 The age of the universe measured back to the big bang depends on the expansion history, which in turn depends on the dark matter and dark energy densities, Ω_M and Ω_Λ. Only the model represented by the *upper curve* includes dark energy. It turns out that the model age is fairly close to the Hubble age, given by $1/H_0$ (BenRG, Wikipedia Public Domain)

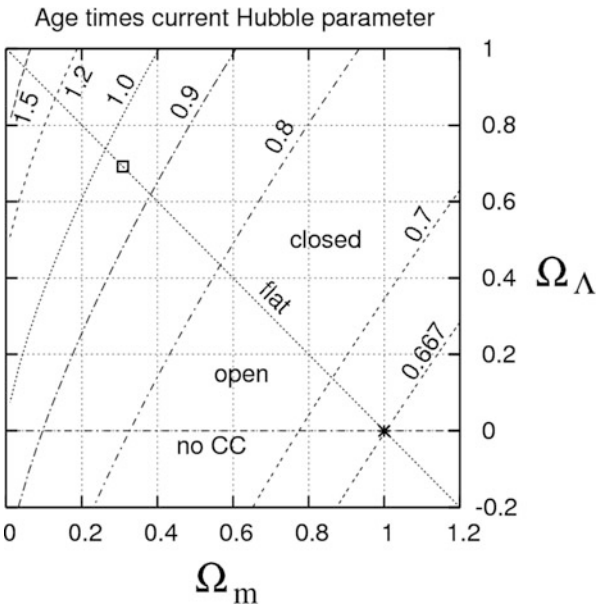


Fig. 1.5 A grid showing the possible world models as a function of the amounts of baryonic plus dark matter (m) and dark energy (Λ). *Curved tracks* show contours of constant age time Hubble parameter (H). The *small box* in the *upper right* shows constraints on the product of the Hubble parameter and age of the universe from observations of the cosmic microwave background by ESA's Planck satellite (Wesino, Wikipedia Creative Commons)

where Ω_M and Ω_Λ are the total matter (dark plus baryonic) and dark energy densities as a fraction of the values required for spacetime to be flat ($\Omega_{\text{total}} = 1$, zero curvature). So if we measure the total matter and dark energy densities and the local expansion rate, H_0, the age t_0 can be calculated in the model (Fig. 1.5).

This is being done using maps of the cosmic microwave background radiation (CMB) from the WMAP and Planck satellites. The CMB is a snapshot of the universe when it cooled enough to no longer be ionized and radiation could travel freely, roughly 380,000 years after the big bang. With a precision of one part in 10^7 for differential temperature measurements, the angular spectrum of the CMB can be fit extremely well with six core parameters of the big bang model. An age of 13.799 Gyr is derived and the 1σ error of 0.021 Gyr deriving from a full Bayesian analysis of uncertainties in all the parameters of the model (Ade et al. 2016).

Measuring the Age Directly

Precision cosmology delivers a surprisingly accurate measurement of the age of the universe. However, it's important to have cross-checks that are rooted in physics and which don't depend on the big bang model.

One direct method involves the radioactive decay of heavy elements where the decay timescale is similar to the expansion timescale, $1/H$. For example, two suitable decay chains are thorium-232 → lead-208 with half-life of 20 Gyr and uranium-238 → lead-206 with half-life of 6.5 Gyr. When the parent element and its

decay product are bound in rock minerals, this method is very reliable, and the Solar System age based on pre-solar grains in meteoritic materials is 4.5682 Gyr with an error of only 400,000 years. But when the heavy elements are distributed in a mixed and evolving system like the Milky Way galaxy, the uncertainty is very much larger. Adding in a formation time for the Milky Way, the uranium chronometer has been used to derive an age of the universe of 14.5 ± 2.5 Gyr (Dauphas 2005).

A second way to measure ages is from stellar models. Since the luminosity of a main sequence star varies as M^3 or M^4, the lifetime varies as $L^{-0.7}$. Therefore, the luminosity of the most massive star remaining on the main sequence gives an upper limit to the age of a cluster containing the star. The mean age of the oldest globular clusters by this method is 11.5 ± 1.3 Gyr (Chaboyer et al. 1998). White dwarfs are remnants of stars with residual heat that slowly radiates into space. The oldest white dwarfs will be the coolest and thus the faintest. Applied to white dwarfs in the globular cluster M4, this method gives an age for the cluster of 12.1 ± 0.9 Gyr (Hansen et al. 2004). As these independent methods all yield answers with ranges that include the age from the big bang model, confidence in the reliability of the CMB-derived age increases.

Time and the Big Bang

Tracing time back toward the big bang, what is the limit to a physically meaningful definition of duration? Early on, the universe is radiation dominated, and matter is coupled and thermalized to the temperature of the CMB radiation. We have:

$$\rho_r \propto g_* T^4, \qquad t = g_*^{-\frac{1}{2}} \left(\frac{T}{1.8 \times 10^{-10}} \right)^{-2}$$

where T is in Kelvin, t is in seconds, and g_* is the effective number of relativistic particles. At very early times, $g_* \approx 100$ in the Standard Model, whereas today only neutrinos are relativistic so $g_* \approx 3$. Projection of time back toward the singularity is tethered in laboratory or accelerator physics back to about a microsecond. In terms of redshift and energy, $T = 2.73$ K $(1 + z)$, so $z \approx T/(1\text{ K})$ and $E \approx T/(3000\text{ eV})$, where T is in Kelvin and E is in electron volts.

The ultimate limit to physical time is the quantum gravity limit, where the de Broglie wavelength in meters is the same order of magnitude as the universe' curvature radius (analogous to the Schwarzschild radius for a black hole), where m is the mass of a particle or black hole and all the physical constants are in metric units:

$$\lambda_{dB} = \frac{2\pi\hbar}{mc} \approx \lambda_S = 2Gm/c^2$$

This sets the Planck scale, which, in the absence of a viable theory to unify quantum phenomena with gravity, is an ultimate limit to knowledge. The Planck scale corresponds to an energy of 10^{19} GeV, a mass of 10^{-5} kilograms, a size of 10^{-35} meters, and a time interval of 10^{-45} s.

Time in the Infinite Universe

Given current measurements of cosmological parameters, the universe will expand at an accelerating rate into the future. If dark energy has the simple properties of Einstein's cosmological constant, this expansion will continue forever. The physics of the far future is somewhat uncertain, depending, for example, on extensions to the Standard Model like supersymmetry, which probably involve proton decay, but the broad outlines can be described.

In the accelerating expansion, an event horizon forms at 16 billion light years, but distant galaxies are steadily removed from view. In about 3–4 Gyr, the Milky Way and M31 interact and merge (Cox and Loeb 2007), and within the combined system, the cycle of star birth and death is broken, and more mass is locked up in stellar remnants and less gas is available to form new stars. All other galaxies leave the horizon 100 Gyr from now. In about 500 Gyr, the lowest-mass main sequence stars exhaust their fuel, and the only new star formation comes from binary systems, where a brown dwarf flickers to life by mass transfer (Adams and Laughlin 1997). The galaxy goes dark, and its mass bifurcates as many stellar remnants "evaporate" into intergalactic space, while others feed the central black hole, which grows 10 billion solar masses. After 10^{25} years, all matter in our galaxy is in the form of isolated stellar remnants and black holes. Expansion stretches photons of the CMB to wavelengths larger than the horizon. For hypothetical observers of the far future, evidence of the big bang will be hard to come by.

Clocks looking back toward the big bang are radioactive decays of heavy atoms and the lifetimes of main sequence stars. Clocks looking forward are pulsars, where the timing for millisecond pulsars is accurate to 10^{-15} s, meaning that technology beats astrophysics for precision but may not win for durability or longevity (Hartnett and Luiten 2011). For a slow pulsar, the spin-down rate is 10^{-15} s per second, and unless they are highly symmetric, they lose energy by gravitational radiation, so pulsars will not be usable as clocks after about 10^{16} years. Radioactive isotopes can be used for 10^{25} years, as tellurium-128 has a double beta decay with a half-life of 2.2×10^{24} years (Audi et al. 2003). These are the longest meaningful timescales in terms of well-established physics.

If baryonic matter decays, then the final phase of cosmic time will be after 10^{30} years, as the supermassive black hole in the galactic center and the smaller stellar remnant black holes throughout the galaxy evaporate (Adams and Laughlin 1997). The power of Hawking radiation increases as the mass decreases, and the evaporation timescale is $t = 8.4 \times 10^{-17} M^3$ s, where M is in kilograms (Page 1976). This timescale is 10^{68} years for a typical stellar mass black hole and 10^{100} years for a

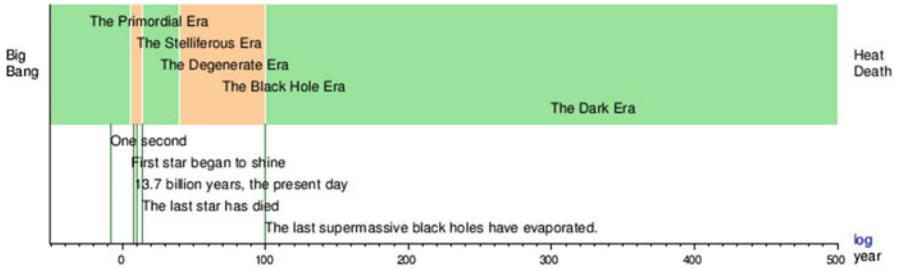


Fig. 1.6 The timeline of the far future of an ever-expanding universe. After the era of stars, there is an era of degenerate matter, then an era of black hole evaporation, and finally a long, slow slide to heat death (Wikipedia Creative Commons)

black hole at the center of a large galaxy. If supersymmetry models that predict proton decay are wrong and baryonic matter is stable, cold fusion by quantum tunneling will turn all remaining stellar remnants into iron-56 in approximately 10^{1500} years. The end point (and final ignominy) of the universe is "heat death" (Fig. 1.6), not because the universe will be hot but because there will be no thermodynamic free energy to sustain processes that can increase entropy, like computation or biology (Dyson 1979).

To bring the discussion full circle, the arrow of time is still governed by increasing entropy even in the bizarrely unfamiliar universe of the far future. In the universe as we experience it, entropy is set by the radiation content of 10^{88} CMB photons, far exceeding the matter content of 10^{80} particles. Black holes have 100 million times more entropy than stars of the same mass; when the universe evolves into black holes, the entropy will grow. The asymptotic state of an endlessly expanding universe is de Sitter space, which has even higher entropy. This progression can be summarized:

$$S_\gamma = \frac{\rho + p}{T} V \approx 10^{88}, \quad S_{BH} = 10^{90} \left(\frac{M_{BH}}{10^6 M_\odot} \right)^2, \quad S_{dS} = \left(\frac{L_{dS}}{M_p} \right)^2 \approx 10^{120}$$

where all quantities apply to the universe as a whole, S is the entropy, V is the volume, T is the temperature, ρ is the density, p is the pressure, M_{BH} is the median mass of supermassive black holes in galaxy centers, and in the final de Sitter phase, the entropy is $A/4G$, where A is the area to the horizon but can also be expressed in terms of L_{dS}, the Lagrangian for de Sitter space, and M_p, the Planck mass. Standard cosmology has no explanation for why the universe began with low entropy, and in most versions of inflation, specific entropy is very low, $10^{10} < S < 10^{15}$, so the situation of special initial conditions is made even worse. This puzzle has convinced cosmologists to explore eternal cosmologies and multiverse scenarios. Time still has not yielded its secrets in the study of the universe.

Conclusions

Cosmology sets the context for any discussion of the nature of time, since the universe is the container within which evolution occurs, and properties of the early universe set the context for the arrow of time, which is not manifested in the time-reversible phenomena of the microscopic world. The universe has been monotonically increasing in size while decreasing in temperature, since the big bang. Cosmic time is measured with respect to fundamental observers, who can apply special relativity, if they are in free fall and not subject to gravity. In the case of general relativity, clocks can be synchronized with an assumption about homogeneity and isotropy, which is in any case required to derive world models. The age of the universe is given roughly by tracing the Hubble expansion back in time and precisely with a multiparameter fit to the angular power spectrum of the cosmic microwave background. The model age of 13.799 billion years agrees with independent checks using radioactive decay rate and stellar physics. The standard model of particle physics sets a limit of a microsecond for talking about verifiable processes in the early universe. Timekeeping in the far future of the infinitely expanding universe can use radioactive decay, pulsars, and, ultimately, the evaporation of black holes.

Acknowledgments The author thanks many of his colleagues at the Steward Observatory and elsewhere for the conversations about time in various astronomical contexts. The author thanks also to the organizers of the "Science of Time" conference for constructing an engaging and interdisciplinary meeting, and the author is grateful to two anonymous referees for helping to clarify and improve this paper. He acknowledges Howard Hughes Medical Institute for the award of an HHMI Professorship and the financial support under grant number 52008138.

References

F.C. Adams, G. Laughlin, A dying universe: the long-term fate and evolution of astrophysical objects. Rev. Mod. Phys. **69**, 337–372 (1997)

P.A.R. Ade et al., Planck 2015 results. XIII. Cosmological parameters, Astronomy and Astrophysics, **594**, A13, 63pp (2016)

G. Audi, A.H. Wapstra, C. Thibault, J. Blachot, The NUBASE evaluation of nuclear and decay properties. Nucl. Phys. A **729**, 3–128 (2003)

J. Avery, *Information Theory and Evolution* (World Scientific, Singapore, 2003)

A. Balbi, Cosmology and time, TM 2012 – The time machine factory, EPJ Web of Conferences. doi:10.1051/epjconf/20135802004 (2013)

G. Belot, Time in classical and relativistic physics, in a *A Companion to the Philosophy of Time*, ed. by A. Bardon, H. Dykes, Oxford, Blackwell, 185–200 (2013)

B. Chaboyer, P. Demarque, P.J. Kernan, L.M. Krauss, The age of the globular clusters in the light of Hipparcos: resolving the age problem? Astrophys. J. **494**, 96–110 (1998)

C.W. Chou, D.B. Hume, T. Rosenbrand, D.J. Wineland, Optical clocks and relativity. Science **329**, 1630–1633 (2010)

R.J. Clowes, A structure in the early universe at $z = 1.3$ that exceeds the homogeneity scale of the R-W concordance cosmology. Mon. Not. R. Astr. Soc. **429**, 2910–2916 (2013)

R.J. Cook, Physical time and physical space in general relativity. Am. J. Phys. **72**, 214–219 (2004)
J.T. Cox, A. Loeb, The collision between the milky way and andromeda. Mon. Not. R. Astr. Soc. **386**, 461–474 (2007)
N. Dauphas, The U/Th production ratio and the age of the milky Way from meteorites and Galactic halo stars. Nature **435**, 1203–1205 (2005)
T. Davis, C.H. Lineweaver, Expanding confusion: common misconceptions of cosmological horizons and the superluminal expansion of the universe. Publ. Astron. Soc. Aust. **21**, 97–109 (2004)
F.J. Dyson, Time without end: physics and biology in an open universe. Rev. Mod. Phys. **51**, 447–460 (1979)
B. Hansen, H. Richer, G. Falhman, P. Stetson, J. Brewer, T. Currie, B. Gibson, R. Ibata, R.M. Rich, M. Shara, HST observations of the white dwarf cooling sequence of M4. Astrophys. J. Suppl. **155**, 551–576 (2004)
J.G. Hartnett, A. Luiten, Comparison of astrophysical and terrestrial frequency standards. Rev. Mod. Phy. **83**, 1–9 (2011)
I. Horvath, J. Hakkila, Z. Bagoly, The largest structure of the universe, defined by gamma-ray bursts, In: 7th Huntsville Gamma Ray Burst Symposium, eConf Proceedings C1304143 (2014)
M. Ibison, On the conformal forms of the Robertson-Walker metric. J. Math. Phys. **48**, 122501–122523 (2007)
A.R. Liddle, *An Introduction to Modern Cosmology* (Wiley, New York, 2003)
L. Maccone, Quantum solution to the arrow-of-time dilemma. Phys. Rev. Lett. **103**, 080401 (2009)
T. Mauldin, *Philosophy of Physics: Space and Time* (Princeton University Press, Princeton, 2013)
C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation* (W. H Freeman, New York, 1973)
D.N. Page, Particle emission rates from a black hole: massless particles from an uncharged, nonrotating black hole. Phy. Rev. D **13**, 198–206 (1976)
J.A. Peacock, *Cosmological physics* (Cambridge University Press, Cambridge, 1998)
M.I. Scrimgeour et al., The WiggleZ dark energy survey: the transition to large-scale cosmic homogeneity. Mon. Not. R. Astron. Soc. **425**, 116–134 (2012)
C. Smeenk, Time in cosmology, in a Companion to the Philosophy of Time, New York, Wiley, doi: 10.1002/9781118522097.ch13 (2013)
D.N. Spergel, The dark side of cosmology: dark matter and dark energy. Science **347**, 1100–1102 (2015)
E.F. Taylor, J.A. Wheeler, *Spacetime Physics. Introduction to Special Relativity* (W.H. Freeman, New York, 1992)

Chapter 2
The Proof of the Pudding

William Andrewes

Abstract In 1775, 2 years after receiving the second half of the longitude reward, John Harrison (1693–1776) published a book, which, among other things, described a pendulum clock that could keep time to 1 s in 100 days. His claim of such unprecedented accuracy for a clock with a pendulum swinging in free air (i.e., not in a vacuum) was met with ridicule both at the time of its publication and for the next two centuries.

This paper describes the remarkable journey of clockmaker Martin Burgess, who set out with a small group of specialists 40 years ago to prove that Harrison's claim was true. A clock build by Burgess according to the principles described by John Harrison was placed on an official trial at the Royal Observatory, Greenwich, in March 2014. After the first 100 days, Burgess' clock was just half-a-second fast. After 2 years of continuous, undisturbed run, its maximum deviation has been 2 s. Currently, it is only 1 s ahead of the atomic time signal. Had there not been such animosity between John Harrison and the Astronomer Royal Nevil Maskelyne, the Royal Observatory at Greenwich could have had a time standard in the eighteenth century that was not realized until the early twentieth century.

Keywords Clock • Greenwich • John Harrison • Longitude • Martin Burgess • Nevil Maskelyne • Precision horology • Time • Timekeeping

This paper describes the remarkable journey of clockmaker Martin Burgess, who set out with a small group of specialists 40 years ago to prove that Harrison's claim was true. A clock build by Burgess according to the principles described by John Harrison was placed on an official trial at the Royal Observatory, Greenwich, in March 2014. After the first 100 days, Burgess' clock was just half-a-second fast. After 2 years of continuous, undisturbed run, its maximum deviation has been 2 s.

W. Andrewes (✉)
Museum Consultant, Sundial Maker, and Historian of Scientific Instruments, Sevres, France
e-mail: wjha@earthlink.net

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_2

Currently, it is only 1 s ahead of the atomic time signal. Had there not been such animosity between John Harrison and the Astronomer Royal Nevil Maskelyne, the Royal Observatory at Greenwich could have had a time standard in the eighteenth century that was not realized until the early twentieth century.

Chapter 3
The Role of Ephemerides from Ptolemy to Kepler

Owen Gingerich

Abstract Celestial timekeeping relied in the first instance on the movements of the stars and planets. The principal systematic positions of planets are recorded in *ephemerides*, which are primarily predictions, not observations. Prior to the invention of printing, ephemerides are extremely rare, which gives lie to the widespread mythology that astronomers before the days of printing were eagerly observing the heavens to add epicycles to improve the accuracy of the tables.

Keywords Ephemeris • Oxyrhynebus • Vienna Papyrus Collection • Alfonsine Tables • John Herschel • Philips Lansbergen • Kepler • J.L.E. Dreyer • Fourier series • Rudolphine Tables

In my personal collection of early astronomy books, I have some volumes of the daily positions of the planets between 1483 (when printing of books was just getting started) and 1626, when Kepler published his third volume of planetary predictions. In contrast, between the fragment of a single leaf from 24 BC to the beginning of printing from movable type in the 1450s, there are currently known only four more or less complete manuscript volumes giving the calculated daily positions of the planets. These four rare volumes are found, one or two each, in the national libraries in Cairo, Paris, and London.

The comparative rarity of such works not only reveals the impact of printing, but, perhaps unexpectedly, it also casts a spotlight on the most pervasive myth in the history of astronomy, a myth I intend to critique later in this essay.

Let us first proceed with a definition of ephemeris: an ephemeris, or in plural, ephemerides, is a tabulation of planetary positions, the most sophisticated outcome of observing the heavens in centuries past. Before planetary positions could be recorded, there had to be a reference frame, such as a particular horizon, or a list of

O. Gingerich (✉)
Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
e-mail: ogingerich@cfa.harvard.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_3

star patterns. With the beginning of writing, records could be kept and patterns noted. The oldest planetary record from which positions can be deduced is the so-called Venus Tablet of Ammisaduqa, which gives on cuneiform tablets the dates when Venus was last seen in the west or first seen before sunrise in the east. These records possibly date from around 1700 BC (although the surviving copies are closer to 1200 BC), and working backward through the ancient Egyptian king lists, Egyptologists deduce that Hammurapi reigned beginning in 1848 BC (Huber 1982). Relatively abundant planetary records survive from around 750 BC, and these helped to anchor period relations up to the time of Tycho Brahe and Kepler. The Venus Tablets are *not* ephemerides, but are very distant cousins.

Let us distinguish four types of related volumes or scrolls. First there are treatises, the detailed "how to" texts, of which Ptolemy's *Almagest* (AD 150) or Copernicus' *De revolutionibus* (AD 1543) are the premier examples. Second are the tables, which expand and tightly organize the numbers from the treatises. Examples are Ptolemy's so-called *Handy Tables* (AD ~150), the Ptolemaic-based Alfonsine Tables introduced in Paris around AD 1320, and the Copernican-based Prutenic Tables (AD 1551). Third are the ephemerides, which are calculated from one or another of the tables. These ephemerides can be very thin, for just 1 year, or very thick, even with as much as half a century of computed positions. Finally there are the light-weight almanacs, generally each for a single year. (Occasionally an ephemeris is titled as an almanac.)

Manuscript ephemerides predating the introduction of printing are extraordinarily rare. The oldest known example, based on a papyrus fragment from the renowned scrap heap in Oxyrhynchus (about 60 miles SSW of Cairo), is dated as 24 BC (Jones 1999). The most recent ephemeris page from Alexander Jones' heroic search, *Astronomical Papyri from Oxyrhynchus*, is from AD 465. The positions of the planets on the fifth-century fragments are generally derived from Ptolemy's *Handy Tables*. A common end-product of these efforts are horoscopes (Thomann 2015a) but the primary purpose appears to be for catarchic astrology, that is, for determining whether particular days were good or bad luck for activities (Thomann 2015b).

For the Islamic period (AD 900–1400), a time for numerous astronomical zijes (texts with astronomical tables), surviving ephemerides are of the utmost rarity, perhaps because they became outdated and were simply discarded. Recently Johannes Thomann stumbled on a significant fragment in the Vienna Papyrus Collection, dating from the year 931–932 AD (Alexander Jones, personal communication). Its physical arrangement shares much in common with later ephemerides: vertical columns for the Sun, Moon, Venus, Mercury, Mars, Jupiter, and Saturn, and the lunar nodes, plus horizontal lines for daily positions. Thomann has taken the lead in discovering fragments of Arabic ephemerides, and so far he has found more than half a dozen discarded examples. The oldest surviving complete Arabic ephemeris was made for the year 1326/27 and is now preserved in the Egyptian National Library (Thomann 2004).

Throughout the Middle Ages and the Islamic period very little was done to improve the positions predicted by the tables. Instead, an effort was undertaken to

make the tables easier to use. Contrary to popular mythology, there was an almost negligible effort to improve the predicted positions derived from Ptolemy's Almagest. There is no paper trail of Islamic astronomers making ephemerides with improved parameters, apart from a more accurate handling of precession (and trepidation). Earlier in the thirteenth century the so-called *Alfonsine Tables* became the most popular basis for calculating planetary positions, and they modified only a single parameter from the Ptolemaic planetary numbers.

On the eve of the first printed ephemerides, two European astronomer/astrologers drew up manuscript ephemerides that still survive, and whose formats anticipate the soon-to-follow printed pages. These formats show a close kinship to the fragments from earlier centuries. One, preserved in the Bibliotheque Nationale in Paris, is for 1426 (see Fig. 3.1). The other, in the British Library, is for 1442–58, and was prepared by the London astrologer Richard Trewythian. Similar to the fragment in the Vienna Papyrus Collection, these full pages have vertical columns for the Sun, Moon, Saturn, Jupiter, Mars, Venus, Mercury, and the lunar nodes, and have horizontal lines for daily positions. Each page contains a month's set of daily

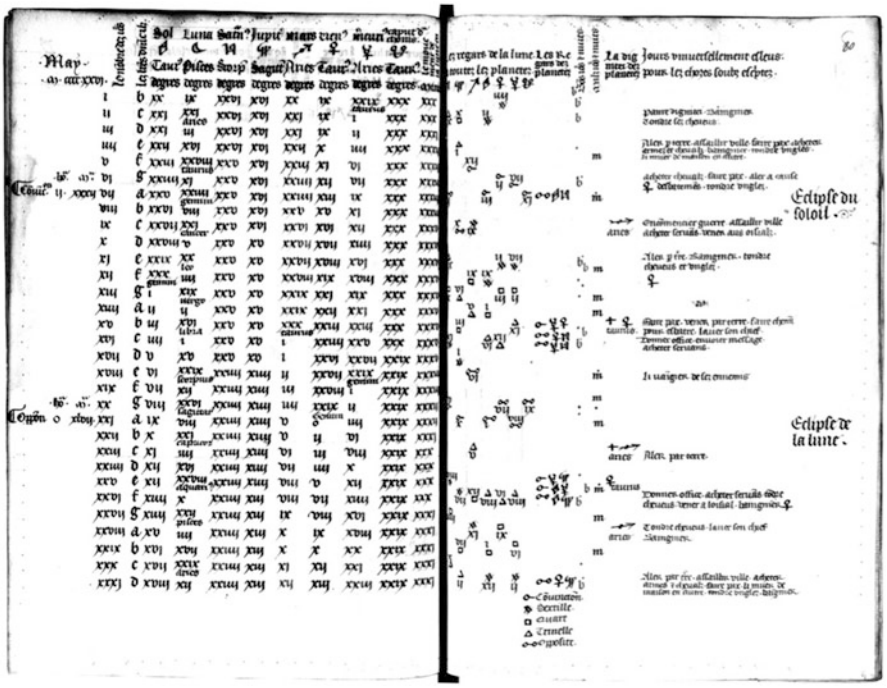


Fig. 3.1 The daily ephemeris for May 1426. To the *left*, the planetary positions day by day. To the *right*, the aspects of the planets with respect to the Moon, then the conjunctions and oppositions with each other. Note the solar eclipse marked in the upper right margin, and the lunar eclipse 2 weeks later. Bibliotheque Nationale, MS Lat 7300 A, fol. 79v-80. By permission of BN

positions. On the facing pages are tabulated alignments between various planets. The positions given in both the BN and BL copies are derived from the widely circulated *Alfonsine Tables*.

With the advent of printing with moveable type, the name of Regiomontanus (1436–1474) comes to the fore. Born in Königsberg in upper Bavaria, young Johannes Muller took the toponym Regiomontanus. He became an acute observer of the heavens, though not necessarily one of the most systematic. His book on trigonometry was a major summary of what was new in a unified subject of trigonometry, though it was not printed until 1533. Nevertheless, Regiomontanus became known as one of the leading mathematicians of the fifteenth century, if not *the* greatest. Besides being an author and translator (from Greek), he became a pioneering printer, particularly for scientific books.

The advent of printing made it reasonable to print run of planetary positions. An individual astrologer might possibly wish to have the birth configurations for a hundred patrons, but this could hardly justify the computation of a thousand or more daily positions for a single, or a few, manuscript copies. The situation changed completely when it became possible to print multiple copies. In the 1470s Regiomontanus calculated 32 years of daily planetary positions using the *Alfonsine Tables,* and in 1474 his shop produced the first printed ephemerides, with daily positions from 1475 through 1506. As Fig. 3.2 shows, it is a masterpiece of

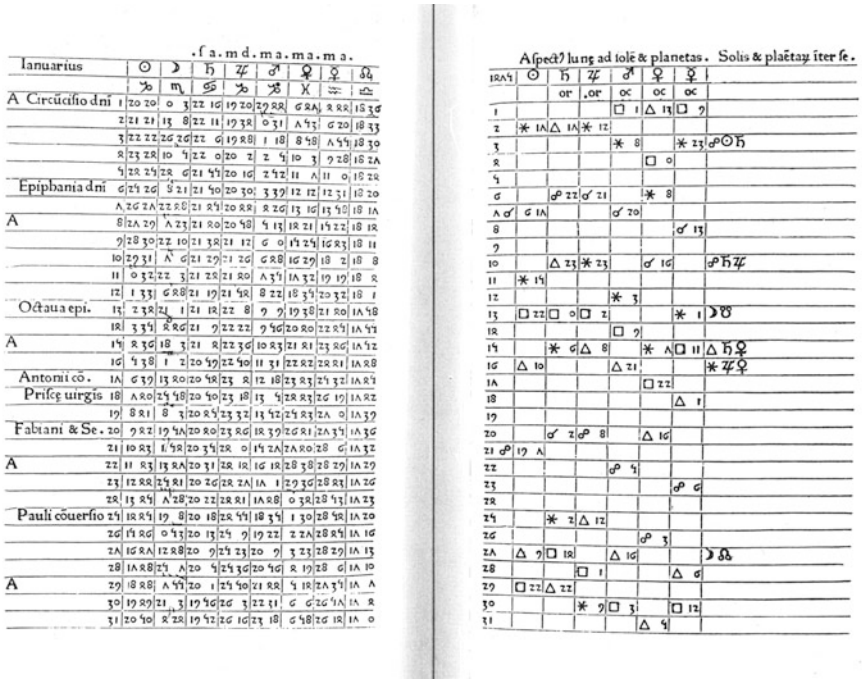


Fig. 3.2 The daily ephemerides for January 1474. Note that the positions (*left*) are given in degrees and minutes within the zodiacal sign shown in the second line. To the *right* are the aspects of the Sun and planets with respect to the Moon. This is the first printed ephemeris, calculated and printed by Johannes Regiomontanus. Facsimile supplied by the author

technical printing, and it closely follows the arrangement of the rare manuscript ephemerides from earlier in the century. In the fifteenth century it was reprinted several times by Erhard Ratdolt, the finest printer of technical materials in the first decades of printing (Fig. 3.3).

This early history of ephemerides has now brought us to the biggest myth in the history of astronomy: that during the Islamic period astronomers noticed the problems with the Ptolemaic predictions, and consequently they added epicycles on epicycles to the Ptolemaic system in repeated attempts to fix the conspicuous errors. The almost complete absence of any relevant manuscript ephemerides suggests that something is totally fake in this modern recreation of how we imagine the study of astronomy should have unfolded. A careful search of possible sources of the mythology suggest that it rose in the nineteenth century, perhaps when the French mathematician Jean B. J. Fourier showed that adding a series of corrections could imitate nearly any function simply by adding more terms. This was hinted at by the astronomer/historian J. L. E. Dreyer, who in 1906 wrote,

> Certainly it would have been possible, as observations multiplied and revealed errors in the theory, to have piled epicycle on epicycle to wipe out these errors, since this is just the same as expressing a function as a series of terms involving sines and cosines of angles proportional to the time elapsed since a certain epoch, which is what astronomers still do (Dreyer 1906).

Much earlier the wide-ranging astronomer John Herschel publicized it in popular textbook *A Preliminary Discourse on the Study of Natural Philosophy* (1831):

> Astronomy, therefore, continued for ages a science of mere record, in which theory had no part, except in so far as it attempted to conciliate the inequalities of the celestial motions with that assumed law of uniform circular revolution which was alone considered consistent with the perfection of the heavenly mechanism. Hence arose an unwieldy, if not self-contradictory, mass of hypothetical motions of Sun, Moon, and planets, in circles, whose centres were carried round in other circles, and these again in others without end,—"cycle on epicycle, orb on orb," till at length, as observation grew more exact, and fresh epicycles were continually added, the absurdity of so cumbrous a mechanism became too palpable to be bourne (Herschel 1831).

Consider the minimum system to represent the motion of Mars. In a first approximation two circles are involved, one representing the earth's orbit, and a second larger circle representing the orbit of Mars. Then the first numerical parameter we need is the comparative sizes of the two circles. Next we need the periods of motion in each circle, and a starting point for each circle, making altogether five parameters. Already several hundred years BC the Babylonian astronomers knew that was not enough. In some parts of the sky Mars moved faster than when it was in the opposite part of the sky. In other words, in another first approximation, there was needed an eccentric displacement, by a specified amount and in a specified direction, so that means two more parameters. This makes a total of seven parameters. And that means we need seven carefully chosen observations to give a fairly decent approximation of Mars' motion.

You can object, for example, to the fact that we have specified no orbital eccentricity for the earth—but after all, in the ancient geocentric system the earth

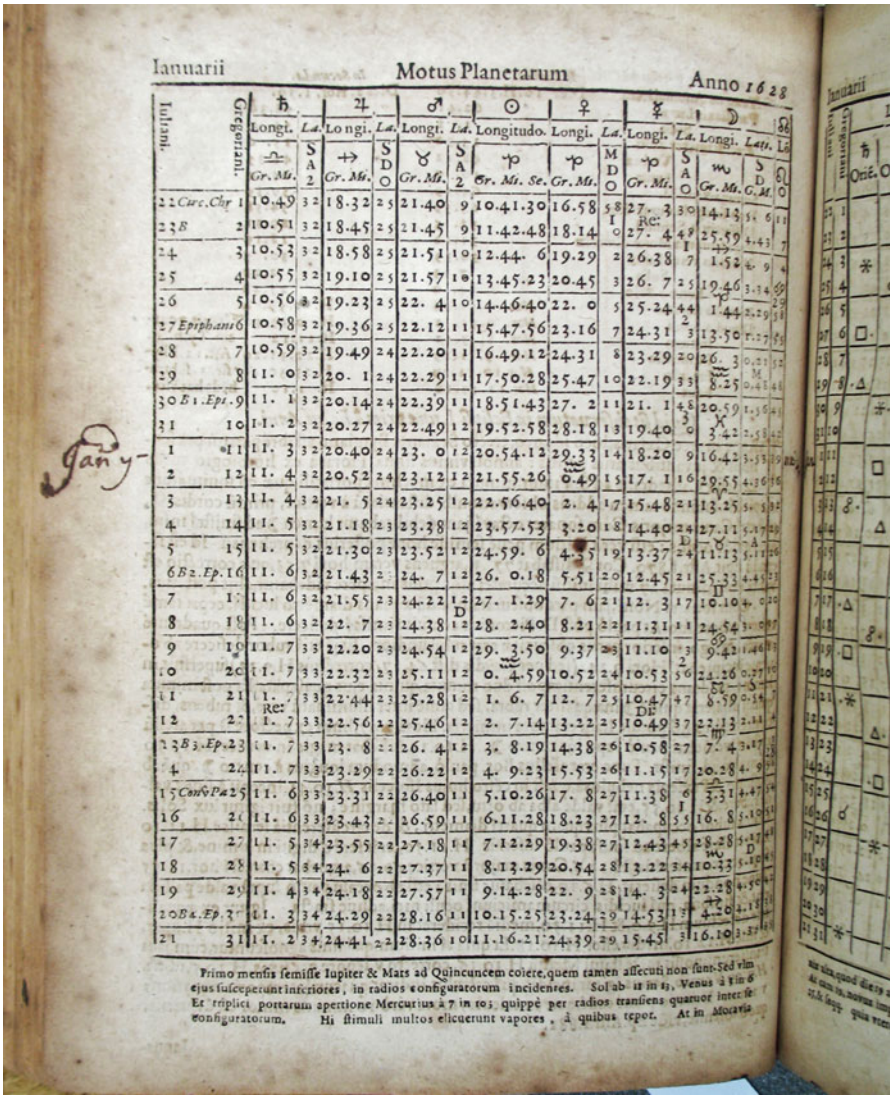


Fig. 3.3 Robert Hooke's sparsely annotated copy of Kepler's Ephemerides, vol. 2, January 1622. Two different calendars were then in use—in Protestant Europe including England the Julian days (*left column*), and in Catholic Europe the Gregorian days (*right column*). Hooke's typical note here flagged the beginning of January, 10 days behind the Gregorian calendar. Collection of the author

was at rest. However, in the geocentric system the Sun is moving, so we could ask for at least, one each, more parameters. On the other hand, these parameters could be folded into the geometry (in this case for Mars). As we will see, this caused the largest glitch in the Ptolemaic system, 5° error for positions of Mars at a critical time every 33 years, but these remained unnoticed until the fifteenth century when Regiomontanus suspected something was wrong with the predicted positions of Mars in 1465, but he lacked the basic observations to correct this anomaly, and thus he simply printed positions derived from the *Alfonsine Tables*.

The first major change in the calculation of tables came with Johannes Kepler's *Rudolphine Tables* of 1628, which improved the accuracy of the predictions by two orders of magnitude (something based on Tycho Brahe's extensive and accurate positions and Kepler's physical insights). Kepler's mentor Michael Maestlin advised him to forget about "all that physics stuff" and use geometry, which was what astronomers were supposed to do. Kepler carried on, using a physics without inertia, but nevertheless calculating the most accurate results that the world had yet seen. But as for the format of his three volumes of ephemerides, the differences were subtle. He now included daily latitudes of the planets as well as their longitudes. Modern calculations show that Kepler did not have the final version of the *Rudolphine Tables* ready for the first volume, which covered 1617–1620. However, the later volumes, for 1621–1636, followed the new tables precisely. Volume 2, which came out in 1630 for the years 1621–1628, carried daily weather records, for those hoping to find correlations between stars and storms.

Shortly thereafter, the Dutch astronomer Philips Lansbergen issued a competing set of tables, much simpler in its structure and rather old-fashioned in their setup. Most astronomers of the 1630s had little way to judge whether Kepler's complex methods of computing planetary positions were really superior to Lansbergen's competing tables, which were much closer in style to the traditional *Alfonsine Tables*. Lansbergen himself compiled a thesaurus or "treasury" of a hundred assorted observations to demonstrate that his methods were better than the world had ever known, though in retrospect he fudged some of the observations. Kepler himself died in 1630, and thus had no opportunity to demonstrate that his ephemerides were in fact significantly superior; a decade later the astronomer Noël Duret demonstrated that indeed Kepler's results were significantly superior.

Although Kepler had printed a thousand copies of the *Rudolphine Tables*, the number of copies surviving today suggests that a few hundreds were simply pulped for the used paper market (Gingerich 2009). But ultimately, despite his mentor's advice, Kepler's physical methods won the day.

Acknowledgements The author wishes to acknowledge the assistance in the early part of this essay from Richard Kremer, Alexander Jones, and Johannes Thomann.

References

J.L.E. Dreyer, *History of the Planetary Systems from Thales to Kepler* (Cambridge University Press, Cambridge, 1906), p. 344

O. Gingerich, in *Kepler versus Lansbergen: On Computing Ephemerides 1632–1662*, ed. by R.L. Kremer, J. Wlodarczyk. Johannes Kepler: From Tübingen to Zagan, Studia Copernicana, vol XLII 2009 (Warsaw, 1970) pp. 113–117

J.F.W. Herschel, *A Prelimnary Discourse on the Study of Natural Philosophy* (London, 1831), p. 266

P.J. Huber, Astronomical Dating of Babylon I and Ur III, Occasional Papers on the Near East, **1**(4), (1982)

A. Jones, *Astronomical Papyri from Oxyrhynchus, American Philosophical Society Memoirs*, 1999, pp. 177–78, plate VII, **2**, 170

J. Thomann, *Article in Chronique d'Egypte LXXXVIII (2013), fasc. 176; see also the photograph of the header in David King*, Synchrony with the Heavens, vol. 1 (Boston, 2004) p. 421

J. Thomann, in *An Arabic Ephemeris for the Year 931–932 CE*, ed. by A. Kaplony, et al., From Bawit to Marw (Leiden, 2015a), p. 115

J. Thomann, in *An Arabic Ephemeris for the Year 931–932 CE, p. 115*, ed. by A. Kaplony, et al., From Bawit to Marw (Leiden, 2015b), p. 116

Chapter 4
How Time Served to Measure the Geographical Position Since Hellenism

Susanne M. Hoffmann

Abstract The orthogonal grid of latitudes and longitudes on a spherical Earth dates back to Ptolemy's book on *Geography* in the second century, and the suggestion to determine the longitude of a traveller dates back to the fourth century. Nevertheless, determining one's own position while travelling remained a problem for two millennia. The solution was technically performed by the Harrison H4 clock in the eighteenth century, but the idea to measure longitudes by time had always been considered a computing exercise. In this paper, we discuss this classical exercise and the question whether and why it had not been applied. Primarily, we recall this example given by different authors and we will recognize differences in their values. Consequently, we will ask about their sources to trace the transfer and transformation of knowledge, since we will finally have to conclude that their descriptions of the measurement were not based on observations but rather that the numerical values were constructed. We further examine the accuracy of the method, investigating if ancient data were sufficient to apply it at all. Summarizing, we can only wonder why scientists for so many centuries praised the method's elegance without collecting the data and applying the method.

Keywords Longitude • Ancient astronomy • Ancient geography • Ptolemy • Hipparchus • Babylonian astronomical diaries • Pliny • Ancient timekeeping

Introduction

"Where am I?" and "where (on the globe) is a certain city?" are important positioning questions for travellers, as well as for astronomers and geographers. This positioning problem has been considered on a spherical Earth since Greek antiquity—probably since the sixth century BC. In these ancient times, both the latitude and the longitude had been determined by measurements of time. The latitude had often been directly given by the ratio of the longest and shortest day and is, therefore, a very simple concept. In contrast, the problem of the longitude had

S.M. Hoffmann (✉)
Excellence Cluster Topoi, Berlin, Germany
e-mail: akademeia@exopla.net

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_4

been solved only in the Modern Era, but the idea to measure the local time of a place to determine its position comes from antiquity.

Longitude: An Unapplied Computing Exercise?

In books of academic astronomy in Greek, Syriac, and Arabic, we can find an example to compute the longitude of places, at least for academic purposes, by comparing local times of astronomical observations. Surprisingly, although all these authors refer to the lunar eclipse in the context of Alexander the Great conquering Babylon in –330, several authors compute with different data in their example. The question, therefore, is where did this schoolbook example and its data come from?

Since they would have realized the differences and possible mistakes if they applied the example, we assume that they never did. The second question is, if their data are useful to determine the longitude, what precision would have been achievable in ancient times?

In this paper, we will first sketch out the classical computational example in all its variations, then discuss possible sources for the input data, and finally estimate the accuracy achievable to determine the longitude with ancient data.

The Method Described by Ptolemy

In his *Geography*, Ptolemy states in Chapter 1.2:

> The surveying component is that which indicates the relative positions of localities solely through measurements of distances; the astronomical component [is that which does the same] by means of the phenomena [obtained] from astronomical sighting and shadow-casting instruments.

What he calls "surveying" is reading data from a map or making a map based on information of *itineraria*. This is a geographer's method. Afterwards, Ptolemy states the important contribution of astronomy:

> Astronomical observation is a self-sufficient thing and less subject to error, while surveying is cruder and incomplete without [astronomical observation].[1]

What Does He Suggest? The technique of measurement is rather simply explained: Two observers measure simultaneously the same quantity. The measured quantity can be either their local time, which is read from a clock or a transiting celestial object. In this case, they would have read the degree of the sky which transits the meridian from an astrolabe-like instrument or so-called

[1] Both quotations: Berggren et al. (2000), 1.2.1, p. 59.

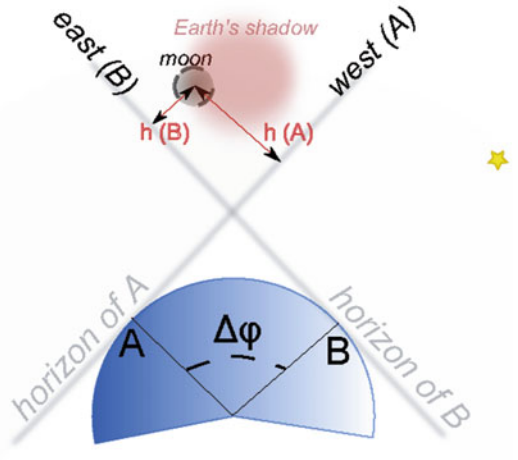


Fig. 4.1 Two places on Earth, *A* and *B*, are separated by the angle Δφ in longitude. Their local horizons are tangent spaces on Earth's surface. The culminating star of *B* is below the horizon of *A* and, hence, invisible in *A*'s place. The Moon entering the shadow is visible for both but at different altitudes with regard to their local horizons

meteoroskopeion.[2] This instrument displaying angles is—in modern terms—an instrument for measuring the local sidereal time. Both methods of measurement—times with clocks and angles with transit instruments—are considered equally convenient.

The problem for ancient observers is to measure something *simultaneously* at two places at a distance. They needed an event which is observable at any place in the *oikumene*[3] and with an appearance independent from time and place on Earth—a so-called absolute event. It suggests that they were looking for celestial events serving as absolute events, and they found lunar eclipses (cf. Fig. 4.1). Ptolemy and his successors suggest observing eclipses "... from which it would have been clear how many equinoctial time units separate the localities to the east or west".[4]

When the Moon enters the shadow of the Earth, it turns darker. This happens at the lunar surface and is, therefore, visible by any observer on Earth simultaneously. Since the event takes place at different local times for two observers, the phenomenon takes place at different height above their local horizon (cf. Fig. 4.1). They can either measure the altitude or azimuth angle of the event or determine the sidereal time of a defined moment (e.g. the beginning of totality) and later compare it with somebody else's observation.[5]

[2]Berggren et al. (2000) point out in a footnote on p. 59 that the *Almagest* only explains "astrolabe", while the "meteoroskopeion" is mentioned and not explained in *Geogr.* 1.3. They add in another footnote that the description of this instrument by Ptolemy is lost but must have existed according to Pappus and Proclus. It was an armillary-like instrument.

[3]The inhabited part of the world (spherical Earth) in Greek is the *oikumene*.

[4]Engl. transl.: Berggren et al. (2000), 1.1.4.

[5]In case of measuring the altitude, of course, the geographical latitude must be taken into account.

As a Variation of This Astronomical Method Ptolemy suggests to determine the angular difference of two places directly by a measurement with a *meteoroskopeion* (μετεωροσκοπίον):[6]

> Using [the meteoroscope] we can easily obtain, among many other extremely useful things, the elevation of the north [celestial] pole at the place of observation on any day or night, and at any hour the direction of the meridian and [the directions] of routes with respect to [the meridian] (. . .). With these [quantities known] we can show right on the meteoroscope the arc in question [of the great circle through the two locations] as well as the [arc] on the equator that the two meridians (if they are distinct) cut off.

Due to our topic, we now want to focus on the suggested time measurement and comparison of the two local times of the observers.

Even though ancient authors are not always specifying the type of eclipses[7] they suggest to use, they all give the same example of one single eclipse serving as standard example for the computation: They always refer to the lunar eclipse of 331 BCE, when Alexander the Great camped in the vicinity of Arbela on his way to conquer Babylon in the famous battle at Gaugamela.

The Classical Calculation Example: Arbela and Carthage

Ptolemy mentions the equivalence of degrees and hours in the *Almagest*, too. In *Almagest*, Book 2, when he argues for the size of the *oikumene*, he states:

> The main proof is that observations of the same eclipse (. . .) by those are the extreme western and the extreme eastern regions of our part of the inhabited world (. . .), never differ by more than twelve hours; and the quarter [of the earth] contains a twelve-hour interval in longitude. . .[8]

He measures the longitudinal difference in hours, but he does not give a reckoning example. In his *Geography*, by contrast, he describes the method and announces this type of measuring longitudes as the method of observing eclipses ". . . from which it would have been clear how many equinoctial time units separate the localities to the east or west"[9]. He considers this method as "a self-sufficient thing and less subject to error, while surveying is cruder and incomplete without",[10] but immediately claims the impossibility of using it in a general way because "most intervals, however, and especially those to the east and west, have been reported in

[6]Berggren et al. (2000), p. 62, *Geogr.* 1.3.

[7]Sometimes we find phrases more generally like "with eclipses" or "with solar and lunar eclipses". The latter variant is probably ill conceived. It cannot work with solar eclipses, since those events are not absolute: Two observers at different places on Earth are hit by the shadow at different times (if they are hit at all). Seen from the Moon, they are not shadowed simultaneously.

[8]Toomer and Ptolemaios (1984), 2.1.

[9]Berggren et al. (2000), *Geogr.* 1.1.4.

[10]Berggren et al. (2000), *Geogr.* 1.2.2.

a cruder manner"[11]. He expresses the opinion that—apart from mathematical reasons—the cause might be that:

> No one bothered to record more lunar eclipses that were observed simultaneously at different localities (such as the one that was seen in Arbela at the fifth hour and at Carthage at the second hour).[12]

Ptolemy certainly was not the inventor of this method. Strabo—living a century earlier—reports it, too, discussing Hipparchus' criticism of Eratosthenes' *Geography*. Hipparchus is said to show correctly that:

> It is impossible for any man, whether layman or scholar, to attain to the requisite knowledge of geography without the determination of the heavenly bodies and of eclipses which have been observed.[13]

The original version of Hipparchus' text to determine the longitude by evaluating lunar eclipses is lost, and only its existence is mentioned in Strabo's *Geography*. Therefore, we do not know whether Hipparchus himself gave an example for the computation and which pair of eclipses or dates of observations he considered—if he applied the method at all.

The third classical author mentioning a pair of eclipses observed simultaneously is Pliny the Elder in his *Natural History*:

> [...] apud Arbilam Magni Alexandri uictoria luna defecisse noctis secunda hora est prodita eademque in Sicilia exoriens.[14]

In English:

> [...] at place of the victory of Alexander the Great a lunar eclipse appears in the second hour of the night and by the same way in Sicily.

Pliny refers to the same eclipse as Ptolemy, but he gives different data, i.e. he compares the eclipse at Arbela not with Carthage but with Sicily, and for Arbela, he states the eclipse at another hour of the night.

While Ptolemy says it took place at the fifth hour, it should have been the second hour of the night according to Pliny. This difference is not explainable by different starting points of counting hours, but theoretically it could be the difference between the beginning and the end of the whole eclipse (not only totality which lasts typically about 1 h). According to the NASA database of eclipses[15] at Sicily and Carthage, the eclipse began before moonrise, and therefore, this cannot be the cause for the difference: If they refer to an observation, they must have considered the end of the eclipse (partial phase) or the beginning or the end of totality.

---

[11] Ptol., *Geogr*. 1.4.2.

[12] Ptol., *Geogr*. 1.4.2.

[13] Strabo and T.E.P (1960), p. 23.

[14] Natural History II, 180 cited from Littre (1877).

[15] http://eclipse.gsfc.nasa.gov/eclipse.html.

What Are the Data Sources for This Standard Example?

Since Pliny and Ptolemy give different times for the lunar eclipse at Arbela, they must have used different sources. But where are eclipses reported? The sources available to them belong to two main categories: the texts of chroniclers who mention it as a sign during a campaign, battle or war and observational data of ancient astronomers.

First, we consider ancient astronomers. To determine differences in local times, they needed two observations, one from Arbela in Mesopotamia and one from another place, either Sicily or Carthage.

In Carthage We Do Not Know Any Astronomers Thus, it is unclear who could have observed there. On the island of Sicily, we know that there was at least Archimedes (died 212 BCE, more than a hundred years after the considered eclipse). Since he is known as a great mathematician and instrument maker, it appears plausible that he was rooted in an important tradition at Sicily although nothing of this came down to us. Thus, we guess that nobody observed at Carthage but only on Sicily, and the two places somehow designate the same observation.

Sicily in this time was confederated with the economically powerful Carthage, and according to later descriptions of sea routes, the island has been important for navigating from Rome to Carthage. It had been an old opinion that early navigators considered Rome and Carthage to be under the same meridian.[16] So, the sources of Pliny and Ptolemy very likely originated at least in the same political entity.

Furthermore, there are reasons for Ptolemy to transform the data he might have found from the school of Syracuse on Sicily to another city, one political reason and a mathematical or educational one: On one hand, it probably appeared to him no big (political) error to declare Carthage to be the origin of the data instead of Sicily. On the other hand, the didactic advantage to do so is derived from the context of his description: When Ptolemy gives the example, he is arguing for the longitudinal extension of the world and, therefore, wants to avoid errors or confusion concerning a difference in latitude. Carthage and Arbela are almost on the same latitude, while Sicily is more north. He might have been led by the intention that the change to another city on the same meridian is a negligible error and, therefore, a price he could easily pay for suppressing complex questions concerning the latitude. As a consequence, the source of Pliny and Ptolemy can be similar or the same, and the only question remaining is, why did they give different hours for an event, which took place almost simultaneously at the two places, and from where did they get the hour concerning Arbela?

Who Could Have Observed at Arbela? Macedonian and Greek historians report the battle of Gaugamela as having taken place at Arbela. Therefore, there must have been several historiographical notions of the event, but we do not know of any astronomer there: neither on Alexander's campaign nor on the Persian place. The village, or caravanserai of Gaugamela, was certainly not famous for astronomy.

[16]Strabo criticizes that in the work of Eratosthenes: Strabo, *Geogr.* 2.1.40.

Thus, when Ptolemy gives precise data of an eclipse in the context of mathematical geography or astronomy, he could not have referred to any astronomer directly in Arbela, but he knew that Arbela is close to Babylon. Since we know that Ptolemy had lists of eclipse tables from Babylon ranging over many centuries, in order to find his source, we should look for this particular eclipse in Babylonian observational reports. Indeed, in the astronomical diaries, we find the message on an eclipse, and a battle is reported 10 days later on the same tablet. However, the layout of these data differs from the data Ptolemy gave:

1 [. . .]⌜ x x ⌝ [. . .]
2 [. . .]⌜ x ⌝ 13 8 ŠÚ⌜ x ⌝ [. . .]
3 [. . .]⌜ GE_6 ⌝ *gab*<−bi>−*šú* ŠÚ 10 UŠ ⌜ GE_6⌝ [GIN . . .] $^{\underline{h}e\text{-}pí}$ MÚL-BABBAR ŠÚ ⌜ GENNA ⌝ [. . .]

and Hunger translates:

> The 13th, moonset to sunrise 8° . . . [. . .] [. . . lunar] eclipse, in its totality covered. 10° night [totality? . . .] (broken) Jupiter set; Saturn [. . .].[17]

How Can This Be Understood?

The first statement "moonset to sunrise 8°" concerns a lunar six interval, which is additional calendar information and not directly important for the eclipse.

Diary entries very often give information, from which direction the eclipse started and after how many UŠ[18] of a partial phase it became total. In this case, there is a gap indicated by "[. . .]" for missing information.

The following statement is typically the duration of totality. Here, we read "10° night [totality? . . .]". Hunger's translation "10° night" is an interpretation of the Babylonian "ten UŠ", which is 10 × 4 min measured by a star or water clock. According to modern calculations, the duration was 58.2 min. Thus, the 10 UŠ (40 min) are within the margin of error (assumed widely).

Since the starting point of the eclipse according to the diaries is not preserved, we are unable to judge whether a diary was the source for Ptolemy or for Pliny.

According to the common Babylonian practice of counting hours from sunset, the missing number in the diary must be something like 40 UŠ, and in Greek notion, this is the *third* hour of the night.[19] This number differs from both the values of Ptolemy and of Pliny and makes it unlikely that one of them got his value from the diary or eclipse tables derived from those as a source.

---

[17]Hunger and Sachs (1988–2014), Vol. 1, p. 176/177, No. −330, lines 2 and 3.

[18]Usually, this term is not translated. We know that one UŠ equals one degree on the equator, equals 4 min in time. UŠ as a distance measure is half a cubit and, therefore, probably a span.

[19]Modern computations show that totality started 2 h 41.4 min after sunset, cf. http://eclipse.gsfc.nasa.gov/eclipse.html or Huber and de Meis (2004), p. 113 or a good planetarium software. This is in the third hour of the night.

Since we proved astronomical texts not to be the source of Ptolemy and Pliny, the remaining proposition is that they found data in texts of chroniclers. Due to the fact that Alexander was celebrated for his contributions to military and imperial history, there were lots of literature about his campaigns in ancient times. There had been novels, poems and reports of contemporary chroniclers as well as Roman historians. Unfortunately, there are only fragments preserved from the chroniclers, who joined Alexander in his wars. In these fragments, no eclipse is mentioned, and we do not know if they would have noticed a lunar eclipse many days before a battle in an ordinary night and found it important enough to tell about. When Chaldean astrology became more important in Roman times, the historians could have written about the eclipse as a sign, and we can assume that Ptolemy and Pliny have read such a story instead of history. Unless there is no information about this eclipse in the texts, which came down to us, we can estimate from other chroniclers' texts, if this type of text might have been a possible source: Looking at typical statements about eclipses in the texts of Greek and Roman historians may give advice.

Herodotus, for instance, states about a historical event "[. . .] when it happened, at an encounter which occurred in the sixth year, that during the battle the day was suddenly turned to night. [. . .] [3]"[20]. This is one of the interesting cases of historiography, because we can try to date kings and events using the mentioned solar eclipse and modern computations.

In many other cases of Greek and Roman historians, we even find no calendar information, but only the information that eclipses (in most cases solar eclipses which occur during the bright day) coincide with a military or political event that the historian reports. Another example is found in Xenophon:

> In the ensuing year, the year in which there was an eclipse of the moon one evening, and the old temple of Athena at Athens was burned. . .[21]

Thucydides is even less precise when he reports:

> In the first days of the next summer there was an eclipse of the sun at the time of new moon, and in the early part of the same month an earthquake.[22]

Eclipses of the Sun are always at the time of the New Moon due to celestial geometry of the event. In sum, we do not learn much about astronomy from the style that chroniclers use, and their information is certainly not suitable for estimating differences in local time of distant cities.

The most precise data are given by Livy:

> On the next night – September 4 – the eclipse took place at the stated hour, and the Roman soldiers thought that Gallus possessed almost divine wisdom. It gave a shock to the Macedonians [. . .] [9].[23]

[20]Herodotus and Godley (2016) Hdt. 1.74.6 cited from http://perseus.uchicago.edu/.

[21]Xen., Hell. 1.6, English translation: Xenophon and Brownson (1918).

[22]Thuc. 4.52, English translation: Thucydides et al. (1910).

[23]Livy, History of Rome, 44.37, Engl. transl.: Livius and Roberts (1912).

But this precision is rare and only found in one Roman text several centuries after Alexander. Although Livy is probably referring to a prediction even he does not state the hour of the night.

Summarizing, in texts of chroniclers, we find only information about the year of the eclipse but nothing more precise. The purposes of timekeeping in common everyday life, which is reported by chroniclers, are not convenient for scientific purposes and especially not for determining longitudes.

If Pliny or Ptolemy looked for data in texts like that discussed above, they did not find data precise enough for their efforts but only stories mentioning the existence of eclipses. If this was the case, the whole example of numbers which they give was probably constructed for education in mathematical thinking and had no source at all.

These results are perplexing, because it means that the method Ptolemy calls the most "elegant" one had not been applied a single time. Ptolemy gives an example and claims a lack of further observational data to be the reason for not applying it. But even the given example does not arise from precise and real data.

Are There Really No Data?

From our point of view, Ptolemy's claim appears to be an exaggeration, because the *Almagest* cites seven lunar eclipses observed by Greek astronomers. If Ptolemy had used the data compiled in his *Almagest* and compared it with Babylonian diary data, he would have had at least seven examples for his so-called superior method, not only one. However, he does not tap the full potential.

The seven *Almagest* examples are (1) a triple of eclipses −200 and −199, which Hipparchus used for computations and are said to be observed at Alexandria.[24] It is not preserved if it was Hipparchus himself who observed the eclipses, but since he used them for computations, we can expect accurate observations, suitable for mathematical astronomy (2). There is a triple of eclipses –133–136 used by Ptolemy himself[25], of which he claims that he had observed them himself. Finally, there is (3) an observation recorded by Heron (62 CE) [26], which Ptolemy cites. It is not sure if all those eclipses really have been seen or only computed by the astronomers cited in the text. Nevertheless, we can state that there are at least seven eclipse observations at Alexandria available to Ptolemy. He could have determined the difference in local time or longitude of Babylon and Alexandria seven times.

Only three of those eclipses had been history when Pliny wrote his *Natural History*, but since Heron's eclipse was contemporary for Pliny, we can question

[24]cf. *Alm*. IV, 11.

[25]cf. Neugebauer (1975), p. 77.

[26]cf. Neugebauer (1975), Vol. 2, p. 780.

why Pliny in Italy did not observe simultaneously with Heron in Alexandria. Of course, there is no answer to that question; we can only state what they did not do. Perhaps they considered this method only as a beautiful example of relations, and a theoretical concept, and did not expect it to be applied, because even after Ptolemy's book with regrets concerning the lack of suitable data, no later astronomer aimed to collect data of pairs of simultaneously observed lunar eclipses. But why?

Is Accuracy the Reason for Not Applying the Method?

Of course the result of all these calculations or observations has an error bar.[27] The error bar of the result, which is calculated as a difference of two observables, is double the observables. The observable in this case is local time at a certain moment or phase of an eclipse. The most accurate measurements of time in antiquity are the measurements of the Babylonian astronomers. They took the time with water clocks, or star clocks (*ziqpu-stars*), and reached accuracies of about 15 min. Let us assume that anybody after Ptolemy—maybe in Alexandria, but that does not matter—continued the Babylonian project and noted every lunar eclipse. This person would have had measured the time during the night with an accuracy of 15 or 25 min, and other observers anywhere in the Roman world would have done the same. Their resulting time difference, after comparing their data, would have had an error of 30–50 min. The difference in local time Δt transforms into a difference in degree of longitude, $\Delta\lambda$, by the angular velocity, ω, of the Earth:

$$\omega = \frac{\Delta\lambda}{\Delta t} = \frac{360^\circ}{24\ \mathrm{h}}$$

This expression is equivalent to

$$\Delta\lambda = \frac{15^\circ}{1\ \mathrm{h}}\Delta t = \frac{15^\circ}{60\ \mathrm{min}}\Delta t$$

Therefore, this error in the computed result of time difference corresponds to a longitudinal error bar of $\Delta\lambda \approx 3.75$.

As an example, we consider the lunar eclipse in −330 at two given places, i.e., at Arbela and at Carthage (according to Ptolemy). The two cities are almost at the same latitude (36°), and according to modern measurements, they differ by $\Delta\lambda \approx 34^\circ$ in longitude. Thus, their difference in local time should be

[27] Although Ptolemy and all astronomers before him did not have the term "error bar", they had a concept of errors in observations (at least since Hipparchus).

$$\Delta t = \frac{34^\circ}{15^\circ/\mathrm{h}} = 2\ \mathrm{h}\ 16\ \mathrm{min}$$

With error bars, the result reads:

$$\Delta t = \frac{34^\circ}{15^\circ/\mathrm{h}} = 2\ \mathrm{h}\ 16\ \mathrm{min} \pm (30 \ldots 40)\ \mathrm{min}$$

Does this result fit to the anciently reported difference of 3 h as the difference of the fifth and the second hour?

According to Ptolemy, yes, because in our standards we have to read his data as $3 \pm \frac{1}{2}$ h or (worst case) as $3 \pm \frac{5}{6}$ h, for the difference in local time. The bigger one of these two error intervals includes the true difference of 2 h 16 min in local time.

Conclusion

The computation of the errors shows that the ancient method of determining a longitude on Earth by the comparison of local times of observers works with an accuracy of 7–10° longitude. Nevertheless, ancient scientific authors had only a handful of eclipses observed simultaneously, and among those, they always chose one famous example to explain the method.

Since only ancient scientist's data, and not the reports of chroniclers, were sufficient to determine the longitude by measuring times, there have been only few data for practical exercises. Ptolemy could have found at least seven eclipses to test and apply his most elegant and most accurate method—but he obviously did not notice this and claimed a lack of data instead. Since additionally no one started expeditions, or called for observations all over the *oikumene*, no ancient cartographer or traveller ever applied this elegant method.

Outlook This state of the art was unchanged during late antiquity. The reckoning exercise has been copied and transformed with some variations by Roman and other astronomers. For instance, we also find it in the books of the Syriac astronomer Severus Sebokht[28] in the seventh century CE but with rough approximations or mistakes in the numbers proving it still bares theory and the input data even constructed for educational purpose.

The elegant method remained an exercise in mathematics until the inventions of John Harrison—the "lone genius who solved the greatest scientific problem of his time"[29] of determining a sailor's position in longitude.

[28]His astronomical opus is discussed in Villey (2014).

[29]This quote is a reminiscence to Dava Sobel's bestselling (hi)story on clock maker Harrison who devoted his life to the development of the world's first portable clock. Such as an eternal flame, it was able to transport the time of a far place with the local observer.

References

L. Berggren, A. Jones, Ptolemaios, *Ptolemy's Geography: An Annotated Translation of the Theoretical Chapters* (Princeton University Press, 2000)

Herodotus, *The Histories*, Herodotus, with an English translation by A. D. Godley (Harvard University Press, Cambridge, MA, 1920). Text provided by Perseus Digital Library. Original version available for viewing and download at http://www.perseus.tufts.edu/hopper/

P. Huber, S. de Meis, *Babylonian Eclipse Observations from 750 BC to 1 BC* (Mimesis, Milano, 2004)

H. Hunger, A. Sachs, *Astronomical Diaries and Related Texts from BABYLONIA*, vols. I–VII (Verlag der Österreichischen Akademie der Wissenschaften, Wien, 1988–2014)

E. Littre, *Collection des Auteurs Latin avec la traduction en français* (Firmin-Didot et c., Paris, 1877). Provided by: http://gallica.bnf.fr

Livius, R. Roberts, *History of Rome*, English Translation by Rev. Canon Roberts. New York, New York. E. P. Dutton and Co. 1912. 1. Livy. History of Rome. English Translation. Rev. Canon Roberts. New York, New York. E. P. Dutton and Co. 1912. 2, provided by http://www.perseus.tufts.edu/hopper/

O. Neugebauer, *A History of Ancient Mathematical Astronomy* (Springer, Berlin-Heidelberg-New York, 1975)

Strabo, H.L. Jones, *The Geography of Strabo* (Loeb Classical Library, Harvard University Press, London & Cambridge, 1960)

Thucydides, *The Peloponnesian War* (J. M. Dent, London; E. P. Dutton, New York, 1910), provided by Perseus Digital Library. Original version available for viewing and download at http://www.perseus.tufts.edu/hopper/

G. Toomer, Ptolemaios, *Ptolemy's Almagest*, translated and annotated by G. J. Toomer (Duckworth, London, 1984)

É. Villey, in *Qennešre et l'astronomie aux VIe et VIIe siècle*, ed. by É. Villey. Les sciences en syriaque (Geuthner, Paris, 2014), pp. 149–190

Xenophon, C. Brownson (ed.), *Xenophon: hellenica* from: Xenophon in Seven Volumes, 1 and 2. Carleton L. Brownson. Harvard University Press, Cambridge, MA; William Heinemann, Ltd., London, vol. 1: 1918; vol. 2: 1921. Text provided by http://www.perseus.tufts.edu

Chapter 5
Changing Times in the Nautical Almanac Over 250 Years

Susan Nelmes

Abstract 250 years ago, in 1766, the first edition of *The Nautical Almanac and Astronomical Ephemeris* was prepared, detailing astronavigational and general astronomical data for the year 1767. Today, fulfilling the same purpose, *The Nautical Almanac*, serving the mariner's celestial navigation needs, and *The Astronomical Almanac*, a worldwide resource for fundamental astronomical data, are still published annually. Since 1981, both books have been jointly published by HM Nautical Almanac Office in the UK and the Nautical Almanac Office of the US Naval Observatory.

Time has played a fundamental role in these publications throughout their 250-year history, from the first edition in 1767 which used Greenwich Apparent Time as the basis of the tabulations, the 1834 edition seeing the change to Greenwich Mean Time, the adjustment of the beginning of the day from noon to midnight implemented for 1925, to the later introduction of Ephemeris Time and Universal Time.

Studying the changes in the timescales used in *The Nautical Almanac and Astronomical Ephemeris* and its successor publications provides an insight into the interface between the ever-changing frontiers of scientific knowledge of time and its practical use by mariners and astronomers.

Keywords HMNAO • Nautical Almanac Office • Almanacs • History • Time • The Nautical Almanac • The Astronomical Almanac

2017 is a significant year for the UK's HM Nautical Almanac Office as it marks the 250th anniversary of our flagship publication, *The Nautical Almanac*. This book has, for two and a half centuries, provided mariners around the globe with the positions of celestial objects allowing them to calculate their position at sea via the techniques of celestial navigation. The creation of such a book and its continued annual production relies heavily on an understanding of time, and to illustrate this, the changes relating to time in the book over its 250-year history can be traced (Fig. 5.1).

S. Nelmes (✉)
HM Nautical Almanac Office, Admiralty Way, Taunton TA1 2DN, UK
e-mail: susan.nelmes@ukho.gov.uk

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_5

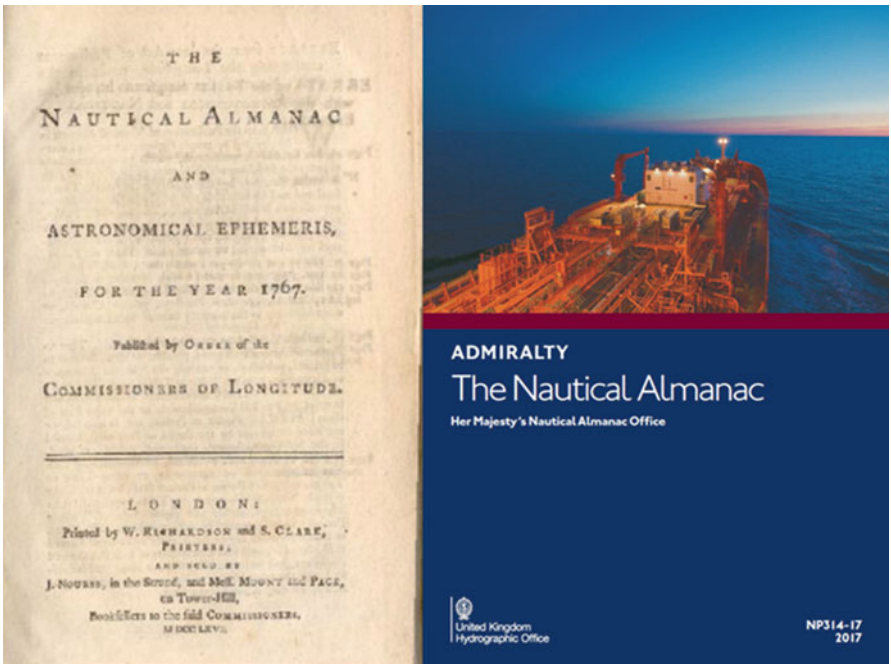


Fig. 5.1 The front covers of the 1767 and 2017 editions of *The Nautical Almanac*

The Nautical Almanac was originally known as *The Nautical Almanac and Astronomical Ephemeris* and first began to be prepared in 1766 after, in the previous year, Nevil Maskelyne, the fifth British Astronomer Royal, presented his suggestion for the calculation and publication of a "Nautical Ephemeris," from which it would be possible to determine longitude at sea using the method of lunar distances. This idea was presented to the Board of Longitude, who, for over 50 years, had offered a prize for such a method or instead a device, for finding one's longitude.

The lunar distances method used the Moon's apparent motion relative to the stars to find a reference time such as the time at Greenwich. The difference between this time and local time on board the ship, which was determined from the position of the Sun, could be converted into a difference of longitude, with 1 h of time difference equating to 15° of longitude.

The time at Greenwich in this case, as the explanation to the first edition of *The Nautical Almanac and Astronomical Ephemeris* describes, was:

> Apparent Time by the Meridian of the Royal Observatory at Greenwich ... deduced immediately from the Sun.

This is nowadays referred to as Greenwich apparent solar time. Using the apparent position of the Sun meant that this was not a uniform timescale due to the variations in the rate of motion due both to changes in the annual motion of the Sun along the ecliptic and of the inclination of the ecliptic to the equator.

To try to counteract some of these problems, the other timescale in general use, although not in *The Nautical Almanac and Astronomical Ephemeris*, was Greenwich Mean Solar Time, using the position of a fictitious point on the mean celestial equator that has a rate of motion virtually equal to the mean rate of the annual motion of the Sun along the ecliptic.

As the explanation to the first edition continues, mean time was the time shown by the:

> Clocks and Watches well regulated at Land.

While this timescale is closer to being uniform, it still does not account for variations in the rate of rotation of the Earth.

Mean time can be found from Greenwich apparent solar time using the equation of time, which has been printed in *The Nautical Almanac and Astronomical Ephemeris* from the first edition and was calculated from the mathematical determination of the fictitious point.

The explanation of the first edition goes on to state that mean time will be of use to astronomers but is of little use to the mariner when using the method of lunar distances. However, it states that if watches, such as those produced by John Harrison, should come into use at sea then:

> The apparent Time deduced from an Altitude of the Sun must be corrected by the Equation of Time and the mean Time found compared with that shown by the Watch, the Difference will be the Longitude in Time from the Meridian by which the Watch was set.

While there were changes in the sources of data, the main pages of *The Nautical Almanac and Astronomical Ephemeris* remained essentially unaltered for almost 70 years. The 1834 almanac saw the first major change with regard to time, along with many other changes which almost doubled the size of the book.

Instead of apparent time as the argument of the ephemerides, Greenwich Mean Time was used instead. This was introduced for the very reason that the explanation in the first edition predicted. It came about, as the 1834 almanac states, because:

> The great perfection of chronometers and their extensive use in sea service have led to the more general use and adoption of mean time; which has now become easy and familiar in many of the most common nautical problems, and equally convenient to all.

The changes introduced in this edition were initiated by recommendations from the Astronomical Society of London, later to become the Royal Astronomical Society, and the French almanac followed this change a year later in 1835 (McCarthy and Seidelmann 2009).

In 1896, the separation of *The Nautical Almanac and Astronomical Ephemeris* into two publications began. This started with a reprinting of the sections most relevant to mariners, rather than astronomers, as a separate volume before this was remodeled as a distinct publication designed especially for the needs of mariners. By 1959, this was known as *The Abridged Nautical Almanac*, while the publication suited to the needs of astronomers was still known as *The Nautical Almanac and Astronomical Ephemeris*. The method of lunar distances was last catered for in the almanacs in 1906 and even this was long after it had stopped being widely used as more modern forms of celestial navigation became popular.

1925 saw the adoption in the almanacs, and elsewhere, of the change of the beginning of the astronomical day from noon to midnight. This was a change from Greenwich Mean Time to Greenwich Mean Civil Time and was designated as such in *The American Ephemeris*. However, since then, Greenwich Mean Time has come to refer to the day starting at midnight, possibly because the British *Nautical Almanac* continued to call the new timescale Greenwich Mean Time (McCarthy and Seidelmann 2009).

This change was first discussed in 1884 at the International Meridian Conference held in Washington where the resolution was passed:

> That this universal day is to be a mean solar day; is to begin for all the 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 is to be counted from zero up to twenty-four hours.

However, it was not until 1917 that the views of astronomers were sought by the council of the Royal Astronomical Society and the American Astronomical Society (Campbell 1918). As this was during the First World War, international consultation may not have led to all astronomers' views being fully represented (Sadler 1967). However, in the UK and the USA, the prevailing, but not unanimous, opinion was that the change should be introduced.

To give time for further consideration and to allow for the change to be implemented in the production of *The Nautical Almanac*, usually produced several years in advance, the change was delayed until January 1, 1925. On this day in the edition of *The Nautical Almanac* for 1925, the date January 1.0 was designated as the same instance as December 31.5, 1924. The ephemerides given in the almanac were given at zero hours both before and after this change and so were given for noon before the change and for midnight afterward. However, the Julian day number, which is a continuous count of days from 4713 BC, still started at noon each day and continues to do so.

As well as changing timescales, HM Nautical Almanac Office has had to decide which times of the day the users of the almanacs would be interested in knowing. The term astronomical twilight had long been established as the time when the Sun had a depression of 18° below the horizon and, naturally, the time when the Sun rose or set has always been important. For the 1937 almanacs, however, the then Superintendent of HM Nautical Almanac Office, Leslie Comrie, introduced two new terms, civil and nautical twilights, corresponding to Sun's depressions of 6 and 12°, respectively (Sadler 2008). These depressions were chosen as they neatly divided the angle between the horizon and the 18° of astronomical twilight. However, they were brought into use because they are useful to the navigator at sea as they roughly correspond with the time when ordinary outdoor activities can no longer be carried out without artificial light for civil twilight and the time when both the brightest stars and the sea horizon are visible for nautical twilight. Their inclusion in the almanacs was discussed with users such as the Dean of the Royal Naval College to ascertain their suitability.

Also in the 1937 edition, times of moonrise and moonset for the southern hemisphere were introduced. Apparently this had the consequence that the Chief

Examiner of Masters and Mates at the Board of Trade complained that he could no longer set one of his standard exam questions (Sadler 2008).

Comrie was also responsible for the formation of the project to predict and reduce occultations of stars by the Moon and planets. To assist with this, a wooden occultation machine built by James McNeile was purchased by the office in 1928 and replaced in 1934 with a metal version (Pratt 2014). This machine enabled the path over the Earth of the Moon's shadow during any occultation to be traced immediately (Comrie 1935). Accurate timings of lunar occultations were of importance in monitoring the motion of the Moon and became the standard method of estimating the value of Delta T, a measure of the difference between a timescale based on the rotation of the Earth (Universal Time) and an idealized uniform timescale at the surface of the Earth (Terrestrial Time), until the second half of the twentieth century.

Recognizing the variability of the rate of rotation of the Earth, the 1950 Paris Conference on the Fundamental Constants of Astronomy introduced the concept of Ephemeris Time, the more uniform timescale based on the orbital motions of astronomical bodies governed by Newtonian dynamics. The name Ephemeris Time was decided by an opinion poll where it competed among other suggestions such as Gravitational Time and the name that came second, Newtonian Time (Sadler 2008).

The period 1954–1960 brought major changes. There was the introduction of Ephemeris Time and all the fundamental changes recommended by the 1952 Rome IAU General Assembly. HM Nautical Almanac Office and the Nautical Almanac Office at the US Naval Observatory worked closely, and from 1960, the UK *Abridged Nautical Almanac* and the American *Nautical Almanac* were unified into *The Nautical Almanac*, using mean time on the Greenwich meridian. Also from 1960, the astronomer-focused publication was jointly prepared by the two offices and re-titled to *The Astronomical Ephemeris*. The fundamental ephemerides of the Sun, Moon, and planets in this publication were based on Ephemeris Time.

The close collaboration between HM Nautical Almanac Office and the US Naval Observatory is still strong, and the publications remain jointly produced and published today under the titles, since 1981, *The Astronomical Almanac* and *The Nautical Almanac*.

Gerald Clemence, the Director of The Nautical Almanac Office at the US Naval Observatory, was a prime mover in the IAU proposal for the use of Ephemeris Time. George Wilkins, later Superintendent of HM Nautical Almanac Office from 1970 to 1989, recalls a conversation between Clemence, Samuel Herrick, a professor at the University of California, and members of HM Nautical Almanac Office in 1952. Wilkins asked if the reform of timescales that included the introduction of the use of Ephemeris Time should not be delayed until atomic time was available as he had heard atomic clocks were close to completion. The reply that he received was that time should continue to be defined by an astronomical phenomenon that would be permanent and not subject to changes in technology (Wilkins 2009).

In 1984, Ephemeris Time in *The Astronomical Almanac* was replaced with the family of timescales known as dynamical time as part of the implementation of the

resolutions decided at the 1976 IAU General Assembly. This family includes Barycentric Dynamical Time and Terrestrial Time which take into account relativistic effects and are still in use today, with their definitions being refined in the 1990s (McCarthy and Seidelmann 2009).

The Nautical Almanac however has continued to use a timescale based on mean time at the meridian of Greenwich up to this day. The term Greenwich Mean Time was solely used to describe this in the British publication of *The Nautical Almanac* until 1985. From then until 1997, the designation Universal Time was used in conjunction with Greenwich Mean Time, and then since 1997, only Universal Time has been used. This is despite the IAU recommending using the name Universal Time in place of Greenwich Mean Time back in 1928 (McCarthy and Seidelmann 2009).

The Universal Time in this case refers to UT1 which, because it is observationally determined by the apparent motions of celestial bodies, is affected by irregularities in the Earth's rate of rotation. George Wilkins recalls that his predecessor as Superintendent of HM Nautical Almanac Office, Donald Sadler, mourned the change in the meaning of Greenwich Mean Time from Greenwich Mean Solar Time to now being aligned, for civil purposes, with Universal Time and used to wear a black tie to mark the occasions, usually on New Year's Eve, when a leap second was introduced (Wilkins 2009).

The use of time in *The Nautical Almanac* is still a topic that we as producers need to be well informed about. One timescale that has been notably absent from this discussion so far has been UTC, Coordinated Universal Time. Celestial navigation is affected by variations in the rotation of the Earth, and UT1 is required. However, the mariner will likely be receiving their time using UTC time signals. Using UTC when carrying out the calculations of celestial navigation using *The Nautical Almanac* can introduce an error of 0.2 min of longitude, which in many cases will be acceptable to the mariner. However, this error is limited by the fact that currently the difference between UT1 and UTC is never more than 0.9 s due to the introduction of leap seconds. If leap seconds were to be discontinued, HM Nautical Almanac Office and US Naval Observatory would have to carefully consider the implications of this decision and find the best way to present nautical almanacs to the mariners that use them.

References

W.W. Campbell, Publ. Astron. Soc. Pac. **30**(178), 358 (1918)
L.J. Comrie, Observatory **58**, 131 (1935)
D.D. McCarthy, P.K. Seidelmann, *Time: from earth rotation to atomic physics*, 1st edn. (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009)
A.R. Pratt, J. Brit. Astron. Assoc. **124**(1), 12 (2014)
D.H. Sadler, Q.J.R. Astron. Soc. **8**, 161 (1967)

D.H. Sadler, A personal history of HM nautical Almanac office. (United Kingdom Hydrographic Office, 2008), http://astro.ukho.gov.uk/nao/history/dhs_gaw/nao_perhist_0802_dhs.pdf. Accessed 5 July 2016

G.A. Wilkins, A personal history of the royal Greenwich observatory at Herstmonceux castle: 1948–1990. (United Kingdom Hydrographic Office, 2009), http://www.lib.cam.ac.uk/deptserv/manuscripts/RGO_history/RGO_GAW-1948-1990_v1.pdf. Accessed 5 July 2016

Chapter 6
Bond Time: The Electric Method of Time Recording

Donald Saff

Abstract From the construction of the first marine chronometer in the United States and the innovation of accurate constant-force clock escapements to the invention of the "Spring Governor" drum chronograph, the Bond family enhanced the accuracy of timekeeping and perfected the distribution of time from the Bond's firm in Boston and the Harvard College Observatory. As a result of their enterprise, combined with the growing network of the telegraph, the "American Method" of transit observation advanced geodesy and positional astronomy. It also provided the mid-nineteenth century burgeoning industry and the New England railroads, through the sale of time, the opportunity to maintain coordinated scheduling. The Bonds were pivotal at the nexus of science and business.

Keywords Marine chronometer • Drum chronograph

From the construction of the first marine chronometer in the United States and the innovation of accurate constant-force clock escapements to the invention of the "Spring Governor" drum chronograph, the Bond family enhanced the accuracy of timekeeping and perfected the distribution of time from the Bond's firm in Boston and the Harvard College Observatory.

D. Saff (✉)
University of South Florida, Tampa, USA
e-mail: donsaff@gmail.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_6

Chapter 7
The Development and Use of the Pilkington and Gibbs Heliochronometer and Sol Horometer

Geoff Parsons

Abstract To meet the continuing requirement for an accurate time standard to calibrate timekeeping instruments, George James Gibbs, later director of the Jeremiah Horrocks Observatory in Preston, invented in 1906 the Pilkington and Gibbs mean time universal Heliochronometer.

Gibbs went into partnership with the businessman William Renard Pilkington who travelled the globe selling Heliochronometers as far as Russia, South America, and Australia, often in regions where the telegraph, and later the radio time signals, were not available.

The Heliochronometer was claimed to provide direct reading mean time within an accuracy of 1 min. To convert solar time to mean time, the instrument used an innovative cam mechanism to apply the equation of time and provided other adjustments for longitude and latitude. The Heliochronometer was produced in six different types, and versions were made for the Northern and Southern Hemispheres. During production, the Heliochronometer underwent several design modifications to improve ease of use and to simplify manufacture.

The partnership entered difficult times and Pilkington developed his own mean time sundial, the Sol Horometer, which looked similar to the Heliochronometer but was sufficiently different to enable a separate patent to be issued in 1911.

Production of both instruments effectively ceased at the outbreak of WWI in 1914. Sales records indicate that only 1000 Heliochronometers were produced, but due to their robust construction, many are known to have survived and are still in use.

The Pilkington and Gibbs Heliochronometer and Sol Horometer are relatively unknown but might be the last sundials in the twentieth century to be used as a practical time standard.

G. Parsons (✉)
North American Sundial Society, British Sundial Society, International Bureau for Weights and Measures, Sevres, France
e-mail: geoffsundial@yahoo.co.uk

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_7

This presentation will explore the purpose, development, and use of these instruments and compare the manufacturing techniques and achievable accuracy. It will place the Pilkington and Gibbs Heliochronometer in its rightful place in the chronological history of timekeeping.

Keywords George James Pilkington • William Renard Gibbs • Mean Time Sundial • Helio Chronometer • Heliochronometer • Sol Horometer • Preston England • Jeremiah Horrocks Observatory

Gibbs went into partnership with the businessman William Renard Pilkington who travelled the globe selling Heliochronometers as far as Russia, South America, and Australia, often in regions where the telegraph, and later the radio time signals, were not available.

Chapter 8
These Are Not Your Mother's Sundials: Or, Time and Astronomy's Authority

Sara J. Schechner

Abstract Drawing upon the exquisite collection of sundials and time-finding instruments at the Adler Planetarium in Chicago—currently being catalogued by the author—this essay offers examples of sundials made of silver, gilt brass, ivory, wood, and stone between 1500 and 1900. They were designed to be portable or fixed, pocket-sized, or monumental, but all did more than tell the time. By critically examining them, we can see the influence of the cultures in which they were made and used. These material objects tell stories of race, empire, labor, religion, fashion, and politics. And by so doing, the sundials exhibit the relationship of time in these concerns.

Keywords Astronomy • Sundials • Time finding instruments • Politics • Religion • Culture • Labor

When most people think about sundials today, they picture a horizontal surface with a triangular sail skimming over a spray of hour lines tipped by Roman numerals. Normally they imagine it cast in bronze and set on a concrete pedestal in a garden. They also believe that this ornamental thing is a better perch for birds than an accurate time finder. Who can blame them? The bulk of sundials sold by garden centers are badly designed, poorly constructed, and capriciously installed. Many simply will not work.

It is hoped that the sundials discussed in this paper will delight and surprise readers who hold dials in such low esteem.[1] They may even rehabilitate dials for them. These sundials were handcrafted by skilled artisans in silver, gilt brass, ivory, wood, and stone between 1500 and 1900. Some were pocket-sized and portable;

[1]This paper draws upon the extraordinary collection of more than 400 sundials and time-finding instruments at the Adler Planetarium in Chicago. The author has been documenting the collection, and a catalogue is forthcoming (Schechner 2017). Particular instruments will be cited using the museum's accession numbers preceded by AP. Please consult Schechner (2017) for additional documentation.

S.J. Schechner (✉)
Collection of Historical Scientific Instruments, Department of the History of Science, Harvard University, Cambridge, Massachusetts, USA
e-mail: schechn@fas.harvard.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_8

others were garden sculpture. Some were humble; others, princely treasure. They exemplify the many forms and styles on which people of different social classes depended to find the time. Their accuracy suited their users' needs, and many could have been employed to set local clocks.

Every one of these sundials is also an astronomical instrument. Each embodies one or more mathematical projections of the celestial sphere and shows the sun's daily motion traced by shadows moving across its various surfaces. They are fitting subjects for a symposium devoted to the science of time. Nonetheless, their production and use as scientific instruments are merely starting points for understanding their historical significance as tangible objects. Like other material culture, sundials are made and employed by human beings and so influenced by cultural milieu and entangled with historical events. To quote Shakespeare, their "nature is subdu'd/To what it workes in, like the Dyer's hand" (Shakespeare 1609, Sonnet 111, ll. 6–7). This means that we can examine the objects in order to discover the impact of sociocultural issues such as race, gender, class, politics, religion, education, and economics (Ulrich et al. 2015; Schechner 2014). Sundials have already been shown to be material evidence of the roles that astronomy has played in people's daily lives (Schechner 2001a, b), and they have been examined as embodiments of cosmological beliefs (Schechner 2007, 2009a). This essay will offer examples of sundials that speak to different aspects of authority.

Authority is a noun of action and force: It is a power to give orders, to make decisions, to enforce, and to control. It implies a recipient upon which action is taken and a vehicle for the action. There are many different ways that sundials as astronomical instruments can be entangled with authority: They can embody edicts of political power or religious order; they can produce and spread astronomical and calendrical information; they can translate and adjudicate between time-telling conventions and can regulate clocks; they can be slaves to fashion or markers of their owners' taste and wealth; they can demonstrate mathematical learning and academic achievement; and they can discipline workers and order lives. The sundials selected to illustrate these points are all part of a remarkable collection at the Adler Planetarium in Chicago.

Political Power

Let us begin with three sundials that evoke political power: an instrument devised by a prince elector, a heraldic crest that found the time, and a travel document carried in the pocket of a lord, diplomat, and military commander.

All that survives of the string-gnomon dial (Fig. 8.1) made in 1547 by Ott-Heinrich, Count Palatine of the Rhine, is an engraved, gilt brass plate, but it bespeaks more than the time. It is a badge of privilege and power. In the center is an elaborate coat of arms emblazoned with four crowned lions, one flanked by buffalo horns and another by wings. The initials "M D Z/O H P" engraved below the coat of arms near 12 o'clock stand for the count's personal motto, *Mit der Zeit* (All in good

Fig. 8.1 A badge of political power and privilege. String-gnomon sundial by Ott-Heinrich, Count Palatine of the Rhine, Palatinate-Neuburg, Southern Germany, 1547. Adler Planetarium, Chicago, M-237

time), and his name and title, *Ott-Heinrich Pfalzgraf*. A member of the Bavarian royal family of Wittelsbach, Ott-Heinrich (1502–1559) was a man to be reckoned with. He was a prince practitioner who made some of his own scientific instruments, a patron of the arts, a book collector, a trustee of Heidelberg University, and a major figure in the Protestant Reformation. From 1505 until his death, he was the Count Palatine of the duchy of Palatinate-Neuburg, a territory in the Holy Roman Empire. In 1556 he added the title of Prince Elector of the Palatinate. Ott-Heinrich, however, was a member of the Schmalkaldic League, a defensive alliance organized by two powerful Lutherans, Philip I, Landgrave of Hesse, and Johann Friedrich I, Elector of Saxony, in response to Charles V's efforts to suppress Protestantism by force in the Catholic Holy Roman Empire. At the time that Ott-Heinrich devised this sundial, his castle at Neuburg an der Donau was under siege by imperial troops. He waited it out. Charles V was victorious in this Schmalkaldic War, but in less than 10 years, he would recognize Lutheranism officially in the Peace of Passau (1552). As Ott-Heinrich's motto affirmed, it was "all in good time" (Schechner 2017, AP: M-237).

The garden dial shown in Fig. 8.2 is a mark of power at the junction of politics and religion. Dating to the seventeenth century and designed for about 44° N, this sundial consists of a slab of polished brown marble that has been sculpted into a

Fig. 8.2 A mark of religious and political authority. Garden dial with heraldic imagery, Italian or South German, seventeenth century. Adler Planetarium, Chicago, M-276

shield with heraldic devices and hour lines for three pin-gnomon dials and one horizontal dial. Lines cut into the wings and tail feathers of the double-headed eagle measure time according to three independent systems of equal hours—the Babylonian (which counted 1–24 from sunrise), the Italian (which counted 1–24 from sunset), and the common (which counted 1–12 twice with starting points at midday and midnight). More will be said about these hour systems below, but here it should be understood that they were appropriate for civil and business purposes. Unequal, planetary hours (which counted 1–12 from sunrise and again from sunset) are incised into the coronet that tops the eagle. This system was used by Catholics to regulate the hours for prayers. That a dial showing unequal hours was placed literally in a crowning position above the heraldic eagle tells us much about the owner's beliefs and alliances. Indeed, one of the two small coats of arms on this instrument belongs to a senior official of the curia or protonotary apostolic (Schechner 2017, AP: M-276).

Political power is also writ large on a gilt brass pocket sundial (Fig. 8.3) made by J. J. Schört of Paris in 1638 for Jacques de Stavay-Molondin (1601–1664), the lord of Molondin, Switzerland, and a French diplomat and military commander. This sundial is a diplomatic pass. The universal equatorial dial is engraved with a gazetteer of 90 cities, providing its owner with the geographic information he needed to set the sundial for any latitude in Europe. The underside of the dial has Stavay-Molondin's coat of arms and a secret code: *Anagramma/Oculus Domini is videt/INDE AMABO DANOS* (Anagram: That eye of God sees, thence I will love the Danes). The meaning of this cipher remains a mystery (Schechner 2017, AP: M-294).

Fig. 8.3 A diplomat's sundial with a secret code. Universal equatorial dial by J. J. Schört, Paris, 1638, owned by Jacques de Stavay, lord of Molondin, Switzerland. Views of the pocket dial with a detail of the underside of the compass box, showing the cipher. Adler Planetarium, Chicago, M-294

Fig. 8.4 A time regulator for a Carthusian monk. Butterfield-type dial by Thomas Haye, Paris, circa 1720, showing St. Bruno praying in the wilderness. Adler Planetarium, Chicago, A-3

Religious Order

The authority of religion is evident in many pocket sundials, and the five discussed here offer diverse examples. Let us consider first a brass, Butterfield-type dial made by Thomas Haye in Paris, circa 1720 (Fig. 8.4). It is a devotional object to humble a monk. On the underside of the compass box, Haye has engraved an image of a man in monastic robes praying on bent knee before a book and a cross in the wilderness. It is St. Bruno. Aspiring to an eremitic life in the desert, Bruno and six companions ventured in 1084 to a rocky, snowy spot in the Chartreuse Mountains of the French

Alps. There Bruno established an ascetic community of contemplative hermits, which became known as the Carthusian Order. The hermitage and "mother house" of the Carthusian Order, Grande Chartreuse, is listed on the sundial's gazetteer (Schechner 2017, AP: A-3). With very little change, Carthusian monks have followed the statutes of their order for more than 900 years. It is said that they live and die by the clock, since their hours are rigorously measured into periods of prayer, spiritual meditation, and manual work, mostly carried out in the solitude of the monks' cells (The Carthusian Order 2016; Merton 1957, pp. 127–144). Table 8.1 shows a typical day for a Carthusian monk. A sundial would be a fitting emblem for them.

Table 8.1 Time discipline typical of a Carthusian monk

11:30 pm	Rise
11:45 pm	Matins of Our Lady
12:15 am	**In church: Night Office (Matins followed by Lauds), lasting 2–3 h depending on the day or the feast**
2:15 or 3:15 am	Lauds of Our Lady. Then back to sleep
6:30 am	Rise
6:45 am	Prime of Our Lady and Canonical Prime. Mental prayer or spiritual reading
7:45 am	To church: Angelus
7:45 am	**In church: Community mass, followed by private masses**
9:00 or 9:30 am	Free time. Mental prayer or *Lectio Divina* (meditative reading of the Bible)
10:00 am	Terce of Our Lady and Canonical Terce
10:45 am	Study and manual work
11:45 am	Angelus, Sext of Our Lady, and Canonical Sext Dinner (always meatless; on Friday, bread and water only) Recreation (reading, working, gardening, enjoying sunshine)
2:00 pm	None of Our Lady and Canonical None. Study
3:15 pm	Manual work
4:00 pm	Vespers of Our Lady
4:15 pm	In church once a week: Office of the Dead
4:45 pm	Free time
5:00 pm	**In church: Canonical Vespers**
5:30 pm	Light evening meal. Time for reading and prayer
7:00 pm	Angelus. Examination of conscience. Canonical Compline and Compline of Our Lady
8:00 pm	Bedtime

This is the typical time schedule for a cloister monk, although there are slight variations between charterhouses, monks, and days. The day is punctuated by the recitation of the Divine Office and the Little Office of Our Lady. Also known as the Liturgy of the Hours and the Little Hours of the Blessed Virgin Mary, they consist of a daily cycle of prayers, psalms, and scriptural readings recited at particular hours. The offices are named Matins, Lauds, Prime, Terce, Sext, None, Vespers, and Compline. The Angelus is a prayer of devotion commemorating the Incarnation and is often announced by a bell. All activities are done in solitude by the monks in their cells, except for communal prayers in the church (shown in bold). (Adapted from the Carthusian Order's official website)

Even in a monastery, daily schedules are subject to shaping by weekly patterns, monthly routines, and the annual cycle of religious festivals and saints' days. For those living outside of religious orders, secular distractions abound. Perpetual calendars help individuals in both communities recall these cycles and are often sundial accessories (Schechner 2014, pp. 52–59). A superb example (Fig. 8.5) is found on the exterior of a silver case that houses an Augsburg-type dial made by Johann Martin between 1671 and 1719 in Augsburg. The case top has a calendar volvelle (a stack of pierced, rotating disks), which aligns days of the week with dates of the month. Apertures reveal the number of days per month, the sun's zodiacal position, the lengths of day and night, and the major saints' days. A lunar volvelle on the case bottom shows lunar phases and the separation of the sun and moon in hours and minutes (Schechner 2017, AP: A-98).

Form and function can also be married in sundials for the devout. Take, for instance, a gilt sundial-reliquary made by the workshop of Ulrich Schniep in Munich, circa 1560 (Fig. 8.6). The cross-shaped hollow box holds the bones of Saints Sabina, Ursula, Ignatius of Antioch, and Walburga, plus a tiny piece of the True Cross clad in gold. The top and bottom faces of the box are engraved with images of a bishop, the Virgin Mary, and assorted Jesuit and Marian symbols. But flip the reliquary over, and the hinged lid becomes the base of an inclining cruciform sundial. The sides and edges of the cross become sundial surfaces and gnomons. The fine quality and imagery suggest that this remarkable sundial probably belonged to a German bishop, who wore it as a pectoral cross or suspended it from his belt (Schechner 2001a, 2017, AP: M-253).

Christianity is not the only religion in which adherents have found time-finding instruments useful in regulating work and prayers. Muslims, for instance, have used sundials and astrolabes to determine the correct times for their daily prayers. During their rituals, Muslims must also orient themselves according to the *qibla*, the direction of Mecca, and, more specifically, that of the *Ka'abah* in Mecca. Given as a compass bearing, the *qibla* varies from place to place on the Earth. Therefore, geographical tables on Muslim instruments give the latitude, longitude, and *qibla* of each city. Specialized devices known as *qibla* indicators attach this information to a magnetic compass with an index that can be set to the right bearing. A sundial may also be included, as in the *qibla* indicator pictured in Fig. 8.7. It was made in England between 1750 and 1800 for the Turkish market (Schechner 2017, AP: A-148).

Religious pilgrims, moreover, must find their way to holy shrines. Figure 8.8 shows an astronomical compendium made of wood in 1513 by Erhard Etzlaub of Nuremberg. In addition to a magnetic compass, the instrument has an incised map of Europe and North Africa complete with latitude scales. Other surfaces have sundials and horary quadrants to find time and available daylight at latitudes 54°, 51°, 49°, 45°, 42°, 37°, 31°, and 24°. Etzlaub created the wooden compendium in order to guide pilgrims to Rome, Jerusalem, and Sinai (Schechner 2001a, 2017, AP: DPW-22).

Fig. 8.5 A perpetual calendar to keep track of religious festivals and lunar phases, engraved on the silver case of an Augsburg-type dial by Johann Martin, Augsburg, 1671–1719. Adler Planetarium, Chicago, A-98

Time-Telling Conventions

In Europe, different countries had different conventions for time telling, and travelers needed to negotiate these changes at the borders. Italians counted their hours 1–24 from sunset, while Bohemians started from sunrise (using a system often called Babylonian hours). In the Nuremberg region, Germans divided the day into two sets of equal hours with the first hour of the day counted at sunrise and the first hour of the night counted at sunset. The French and English, however, preferred the system of common hours in which the day was divided into two groups of 12 equal hours starting at noon and midnight. By displaying plural hour systems, portable dials enabled their users to convert between them (Schechner 2014, Chap. 5).

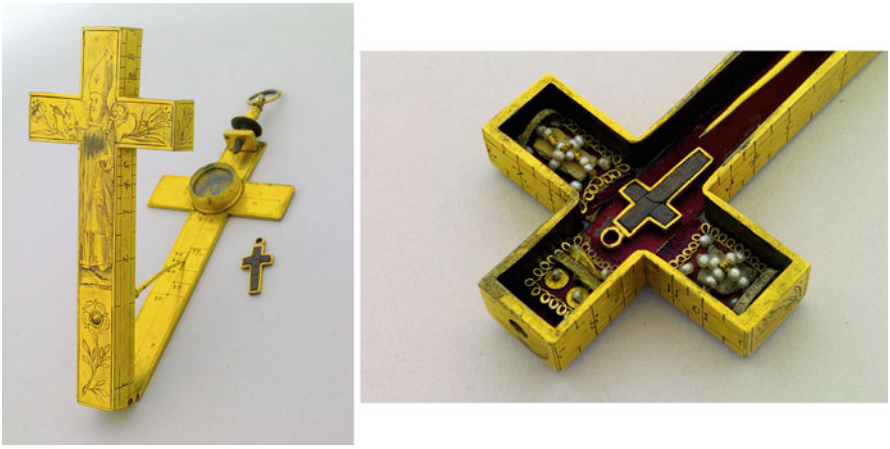


Fig. 8.6 A bishop's reliquary in the form of a cruciform sundial, made in the workshop of Ulrich Schniep, Munich, circa 1560. Adler Planetarium, Chicago, M-253

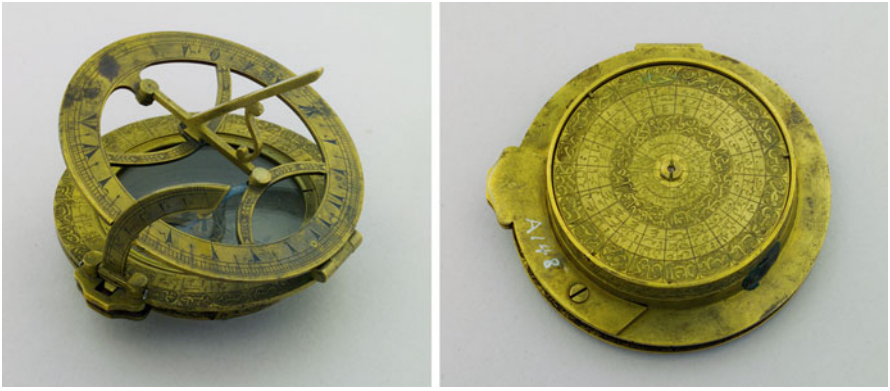


Fig. 8.7 A means to find the direction of Mecca and hours of prayer. Qibla indicator and inclining dial, probably made in England for the Turkish market between 1750 and 1800. Adler Planetarium, Chicago, A-148

Figure 8.9 illustrates the bundling of hour systems on an ivory diptych dial by Thomas Tucher of Nuremberg, circa 1610–1620. A string, now missing, once stretched between the inner faces of the lower and upper leaves and served as a gnomon for four sundials. These all read common hours. Four pin-gnomon dials on the vertical and horizontal faces show Italian hours, Babylonian hours, and the length of daylight at different times of the year. In the daytime, a person wanting to read Nuremberg hours can use the Babylonian-hour dial, which is labeled "Nirenperger Uhr." At night, the Italian-hour dial will serve him by moonlight if he corrects the shadow reading for the number of hours that the moon is behind the sun. To get this right, he would take advantage of the lunar volvelle on the outer

Fig. 8.8 A guide for pilgrims on the way to Rome, Jerusalem, or Sinai. Wooden compendium by Erhard Etzlaub, Nuremberg, 1513. Adler Planetarium, Chicago, DPW-22

face of the lower leaf. Note that the Italian-hour dials are labeled "Welsche Uhr" (foreign hours), highlighting the perspective of the maker in Nuremberg (Schechner 2017, AP: T-10).

Another example is the ivory diptych by Johann Gebhart, Nuremberg, 1546 (Fig. 8.10). This instrument has an equatorial dial on its top leaf showing common hours. In its center is a lunar volvelle. Instructions incised into the ivory remind the user how to convert the sundial into a moon dial. The lower leaf has a pin-gnomon dial whose crisscrossing lines indicate both Italian and Babylonian hours (Schechner 2017, AP: DPW-21).

Sundials like these examples were treaties on time zones.

Slaves to Fashion

Like stylish clothing and cooking, sundial designs followed fashion trends, particularly in major cities of production. It was not enough to simply find the time. Consumers wanted to show off their good taste when they pulled a sundial from their purse or pocket. Each city had its special style. In late-seventeenth- and eighteenth-century London, compass dials were popular (Fig. 8.11), whereas in Paris, Butterfield-type dials with their cute bird-gnomons were all the rage (Figs. 8.4 and 8.12). In Augsburg, individuals purchased universal equatorial dials in a form that became synonymous with the city (Figs. 8.5 and 8.13), while

Fig. 8.9 Plural hour systems for use by travelers. Ivory diptych dial by Thomas Tucher, Nuremberg, 1610–1620. Adler Planetarium, Chicago, T-10

the Nuremberg compass makers produced diptych dials (Figs. 8.9 and 8.10). Dials purchased by the wealthy were fashioned of silver, gilt brass, or ivory. Those of lesser means bought cruder instruments of brass or wood (Schechner 2001a, 2017, AP: A-95, A-243, W-237).

Luxurious goods also existed. Exquisite crescent dials of silver and gilt brass were imported from Augsburg to London. An example by Johann Martin (Fig. 8.14) is signed by Masig, an import agent in the early eighteenth century (Schechner 2017, AP: M-301). Butterfield-type dials by the inventor, Michael Butterfield, were so desirable that they were counterfeited (Schechner 2001a, 2009b, 2017, AP: N-5). Instruments such as the polyhedral dial fashioned by Ulrich Schniep of Munich in 1577 (Fig. 8.15) had the status of *haute couture*. Shaped as a gilt tower, this windowsill instrument has sundials on every face, a magnetic compass on the underside of the domed roof, and a level within its chamber. Schniep had many noble clients, including Duke Albrecht and Emperor Maximilian II. This polyhedral dial was made for a member of the von Stadion family of Augsburg, whose coat of arms appears on the instrument (Schechner 2017, AP: M-304).

Fig. 8.10 A treaty on time zones. Ivory diptych sun and moon dial by Johan Gebhart, Nuremberg, 1546. Adler Planetarium, Chicago, DPW-21

Fig. 8.11 London fashion. Compass dial, Samuel Rolfe, London, 1742. Adler Planetarium, Chicago, W-237

Fig. 8.12 Paris fashion. Butterfield-type dial, Michael Butterfield, Paris, 1681–1684. Adler Planetarium, Chicago, A-243

Fig. 8.13 Augsburg fashion. Augsburg-type dial, Johann Mathias Willebrand, Augsburg, 1700–1725. Adler Planetarium, Chicago, A-95

Authority of Learning

Complexity in polyhedral and multiple-faced instruments such as the Schniep tower-shaped sundial offered owners the opportunity to show off not only their breeding and good taste but also their learning. Such dials also establish their makers' virtuosity. As the sun travels across the sky, the sundials on one face hand off the job of finding time to those on the next.

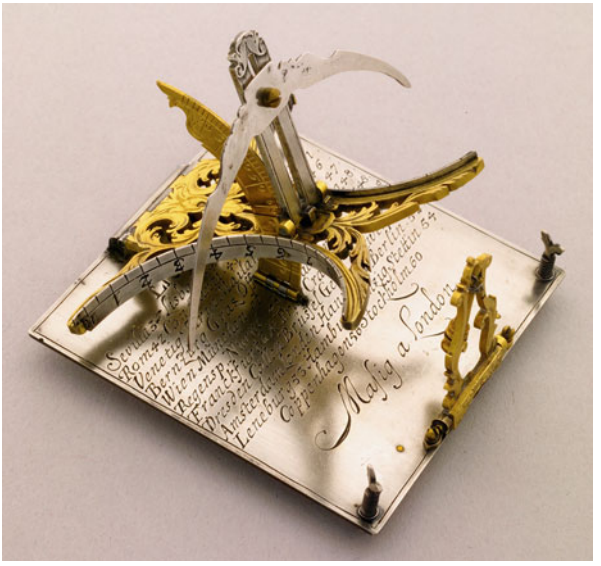


Fig. 8.14 A stylish import. Crescent dial by Johan Martin, Augsburg, imported to London by Masig in the early eighteenth century. Adler Planetarium, Chicago, M-301

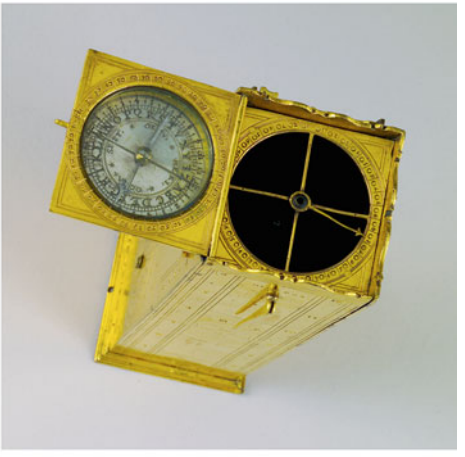


Fig. 8.15 Haute couture. Tower-shaped polyhedral dial by Ulrich Schniep, Munich, 1577, for a member of the von Stadion family of Augsburg. Adler Planetarium, Chicago, M-304

The silver polyhedral dial, built in Germany about 1620 (Fig. 8.16), is deliberately overengineered to be a spectacle. Based upon a design published by Oronce Fine in his *Protomathesis* (Paris, 1532) and again in *De solaribus horologiis et quadrantibus* (Paris, 1560), this object is a conglomeration of planes, cavities, rings, and spheres. It has 29 distinct dials on its various surfaces. This is not a sundial; it is a dissertation in mathematical astronomy! (Schechner 2001a, b; Schechner 2017, AP: M-324).

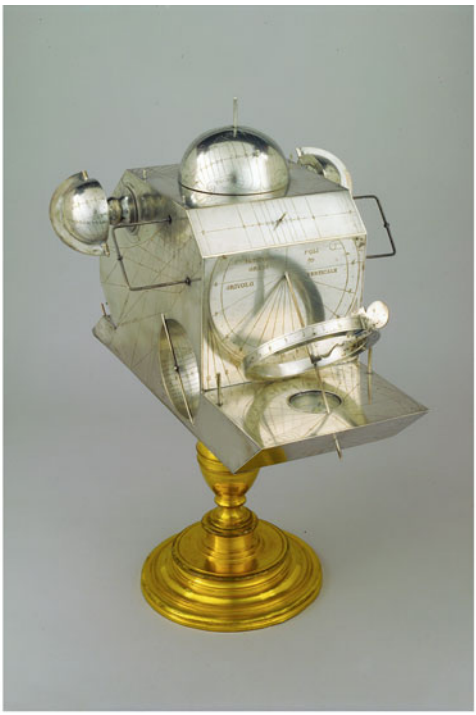
Labor Relationships

Modern societies are characterized by a division of labor, a hierarchical structure of workers, and schedules for particular types of work. Sundials have exhibited all of these labor relationships.

First and foremost, as time-finding devices, sundials have always encouraged time discipline (Schechner 2001a, 2014). Even the simplest ones show the passage of time and can admonish the onlooker to use his or her time more wisely. But sundials that provide information on the lengths of days versus nights during the course of the year have a greater capacity for time management, since they enable the user to plan ahead. For example, the table-sized dial by Ulrich Schniep of Munich, 1572 (Fig. 8.17), has declination arcs showing the sun's place in the ecliptic. These are divided into zodiacal signs in the order of day duration increasing or decreasing. Along the meridian line of the hour plate are numbers indicating the hours of daylight depending on the sun's position (Schechner 2017, AP: M-243). Such a sundial would have made a very wealthy merchant, banker, or senior administrative official mindful of his business.

Sundials could also be attached to the tools of people who worked with their hands as well as heads. A late-seventeenth-century example is the sundial mounted over a surveyor's plane-table compass (Fig. 8.18). This special type of magnetic compass has a square base with straight edges. A surveyor employs it on top of a portable flat surface, called a plane table, when drawing charts and maps in the field. A sundial accessory would have helped the surveyor keep track of the time. This example was made by Thomas Haye, Paris, 1690–1700 (Schechner 2017, AP: DPW-2).

As manufactured goods, sundial production followed the rules of local guilds, which concerned the economic and training relationships between masters, journeymen, and apprentices. Sundials also expressed relationships between patrons and clients with the craftsmen that served them. To this end, the carved limestone, garden dial (Fig. 8.19) made in 1719 by Franz Xaver Josef Bovius in Eichstätt, a town in upper Bavaria, is not just a sundial; it's a job application. Bovius was a Catholic priest, who is known to have made sundials for many nobles in his chapter. This particular dial was made for Casimir Ferdinand Adolf von Waldbott of Bassenheim, a nobleman and choir bishop in Mainz and Trier. A dedication to Waldbott, which includes all his titles and distinctions, is carved into a banner



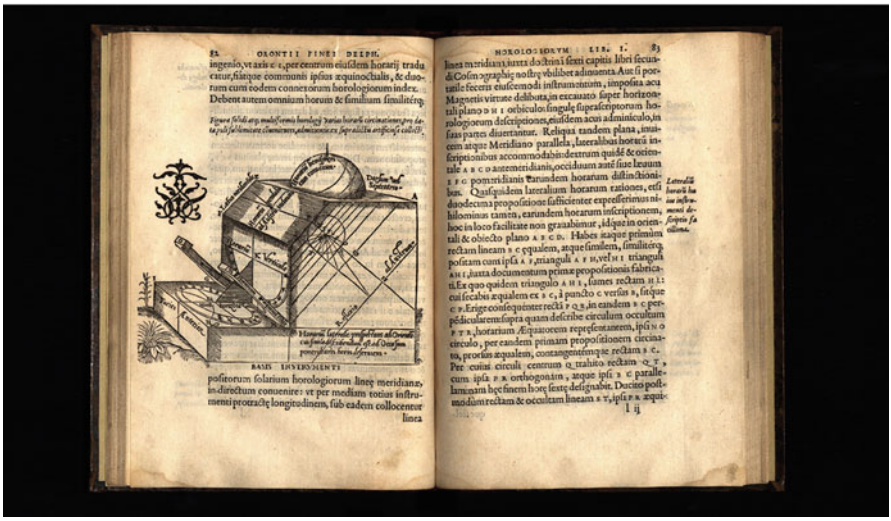


Fig. 8.16 A mathematical thesis. A German polyhedral dial, circa 1620, based on a design published by Oronce Fine in *De solaribus horologiis, et quadrantibus* (Paris, 1560). Adler Planetarium, Chicago, M-324

Fig. 8.17 A work planner. String-gnomon dial by Ulrich Schniep, Munich, 1572, showing the hours of daylight available each month. Adler Planetarium, Chicago, M-243

above Waldbott's coat of arms. The arms are also engraved four times onto the brass gnomon. A central element of Waldbott's heraldic crest is the swan, and Father Bovius has carved swan motifs no less than 35 times in emblems that encircle the sundial. Mottos with the emblems suggest Waldbott's attributes (fortitude, prudence, temperance, justice, brightness unimpeded, trusting, not weakened, touched but not shaped, looking to the heavens, following God, and so forth). Others are fawning affirmations such as "Scribe Beati" (Write of the happy), "Latior in Senio" (I am brought forth in old age), and "Excelsa Peto" (I demand the heights). Three chronograms[2] on the sundial also reference the red and white colors of Waldbott's coat of arms, and one enjoins, "IVgIter soL nobIs ConserVabIt aDoLphVM" (The sun will perpetually preserve our Adolf). By copying and rearranging the capital letters, we get MDCLLVVVIIII, the year 1719 in which

[2]Chronograms are sentences or phrases in which specific letters are interpreted as Roman numerals, which stand for a particular date when pulled out of the sentence and rearranged.

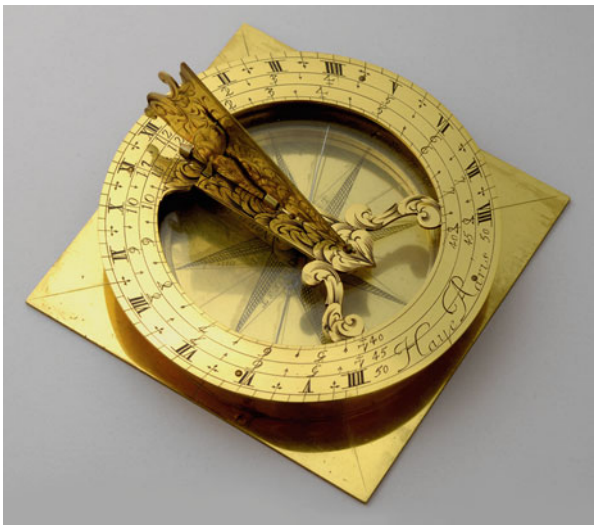


Fig. 8.18 A task master. Surveyor's plane-table compass with a built-in sundial, by Thomas Haye, Paris, 1690–1700. Adler Planetarium, Chicago, DPW-2

the dial was made. No sundial has more flattery of its patron than this. Father Bovius's signature reminds the recipient that he has been examined, approved a candidate of canon law, and served as priest in the diocese of Eichstatt (Schechner 2017, AP: M-286).

A very fine sundial by Christoph Schissler, Augsburg, circa 1562 (Fig. 8.20), offers another perspective on labor relations. The gnomon (now missing) was a string attached to a staff held by the turbaned figure of a Moor. The exotic black African represents a slave at a royal court or in the household of a European nobleman. He symbolizes his owner's power at home and abroad. He is at his master's beck and call at all hours (Schechner 2017, AP: M-240).

Empire

The mention of slave labor brings us to the topic of imperialism and how sundials served colonial administrators and military officers. Figures 8.21 and 8.22 show two examples of brass pocket sundials carried by French military officers across Canada and down the Mississippi in the mid-eighteenth century. Both were made in Paris by Pierre le Maire II between 1730 and 1760 and are among six extant examples (Schechner 2016). The underside of each carries a gazetteer for New France listing colonial capitals (e.g., Quebec, Port Royal, and New Orleans), forts large and small (e.g., the Fortress of Louisbourg and Fort Louis), fur trading posts (e.g., Green Bay, Montreal, Fort St. Anne, and Michipicoten), and Native American tribes (e.g., the Iroquois, Abenaki, Wabash, Missouri, Illinois, Quapaw, Natchez, Pensacola, and others). The sundials are evidence of the transmission of cartographic and ethnographic knowledge gathered by missionaries, explorers, soldiers,

Fig. 8.19 A job application. Garden dial designed to impress the recipient and include the maker's resume, carved by Father Franz Xaver Josef Bovius of Eichstätt, Germany, 1719, for Casimir Ferdinand Adolf von Waldbott of Bassenheim, a choir bishop in Mainz and Trier. Adler Planetarium, Chicago, M-286

and colonial merchants. These wilderness sites were bitterly contested by France and England in numerous battles and proxy wars involving Indian tribes. The sundials, however, made these Indian nations seem smaller by confining their ranges to very specific latitudes. It is also sobering to see the city of La Rochelle on one sundial. This French port was central to France's triangle trade, which transported slaves from Africa to the West Indies and Louisiana, and sugar from plantations there back to France (Schechner 2016, AP: T-58, W-57).

Similar sundials were made for use in New Spain (Fig. 8.23). An equatorial dial by Juan Andres, La Paz, 1699 (A-263), is made from the silver being mined nearby and exported to Spain. Its gazetteer lists locations in the northern part of the Viceroyalty of Peru, including Potosí, known for its great silver lodes; Lima, the capital of the viceroyalty and hub for the export of silver and import of

Fig. 8.20 Slave labor. A black African Moor once held the string-gnomon on this lavish sundial by Christoph Schissler the Elder, Augsburg, circa 1562. Adler Planetarium, Chicago, M-240

Fig. 8.21 A colonial administrator. Butterfield-type sundial made in Paris by Pierre le Maire II, 1730–1760, for use in New France. The gazetteer lists fur trading posts, forts, and Native American tribes across Canada. Adler Planetarium, Chicago, T-58

manufactured goods; Cusco and Quito, producers of cotton and woolen textiles; and Arequipa and Nazca, producers of wine. The gazetteer also includes indigenous people. The only ornamentation on this sundial is a simple *globus cruciger* (cross-on-orb), symbolizing the dominion of Christianity over the globe. The decoration

Fig. 8.22 A dial for conquerors. Butterfield-type dial by Pierre le Maire II, Paris, circa 1750, for use in French possessions down the Mississippi River and its tributaries and east along the Gulf of Mexico. The gazetteer lists indigenous peoples to be governed: the Wabash, Missouri, Illinois, Arkansas (Quapaw), Natchez, Pensacola, and Tohomé. La Rochelle, a coastal city in France, is also included. The port was involved in France's trade of slaves from Africa and sugar from plantations in the West Indies and Louisiana. Adler Planetarium, Chicago, W-57

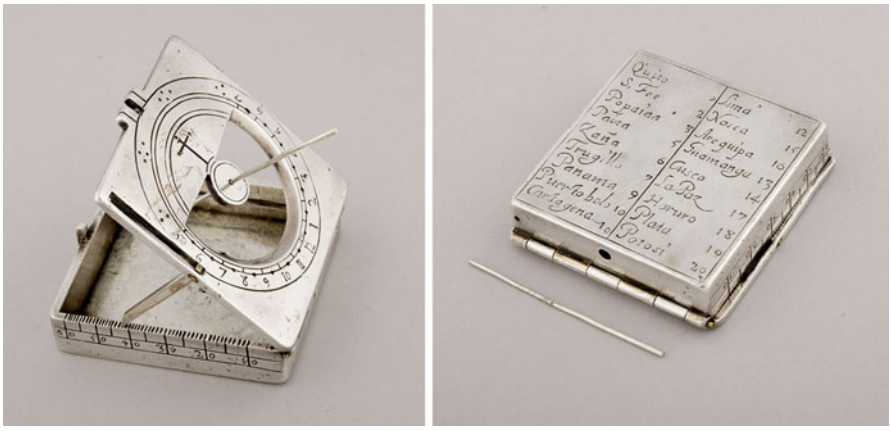


Fig. 8.23 On a mission in New Spain to import Catholicism and export silver. Equatorial sundial by Juan Andres, La Paz, Bolivia, 1699. Adler Planetarium, Chicago, A-263

suggests that this sundial was made for missionaries, but the locations were also mining, textile, and agricultural centers with European- and colonial-born Spanish administrators and landowners, who oversaw a labor force of enslaved Indians and some Africans.

Clock Regulation

The last section of this paper examines the tug-of-war between apparent solar time and mean time as represented by the material culture. Clocks play such an outsized role in modern life that it is easy to forget that sundials had the upper hand in accuracy and affordability well into the nineteenth century. Sundials capable of distinguishing minutes were produced a century before minute hands became common on clocks. Clocks, moreover, could only keep mean time from an arbitrary starting point, whereas sundials and other time-finding instruments found the time directly from the sun or stars. Therefore, every clock user needed a good time-finding instrument to set his or her timekeeper. One early solution was to buy a clock or watch with a sundial attached. Another took advantage of specialized sundials that were noon markers. The most popular noon markers throughout the nineteenth century were cannon dials set up in public parks (Figs. 8.24 and 8.25). Sunlight focused by a lens on the touch hole of a small cannon announced noon with a blast that delighted onlookers (Schechner 2017, AP: W-104).

As watches and clocks became more accurate, people cared about the discrepancy between apparent solar time (read by a sundial) and mean solar time (measured out by a clock) and began to favor mean time. To convert between the two, tables for the equation of time were printed in almanacs and on paper disks that were pasted inside the cases of pocket watches and sundials. With the familiar phrases, "Watch Fast" and "Watch Slow," the tables informed users whether to add or subtract minutes from the time read off their sundials in order to set their clocks to mean time. By the late eighteenth century, those who found this too tedious could purchase a heliochronometer, a sundial that found mean solar time directly. Johann Jakob Sauter II of Stockholm, 1802–1811, produced the fine example shown in Fig. 8.26, which mechanically corrects for the equation of time (Schechner 2017, AP: M-302).

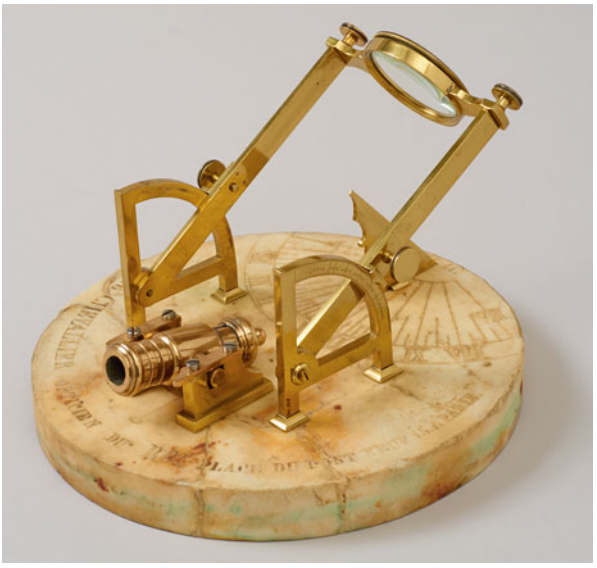


Fig. 8.24 An acoustic sundial to regulate clocks. Cannon dial by Jean-Gabriel-Augustin Chevallier, Paris, circa 1820. Adler Planetarium, Chicago, W-104

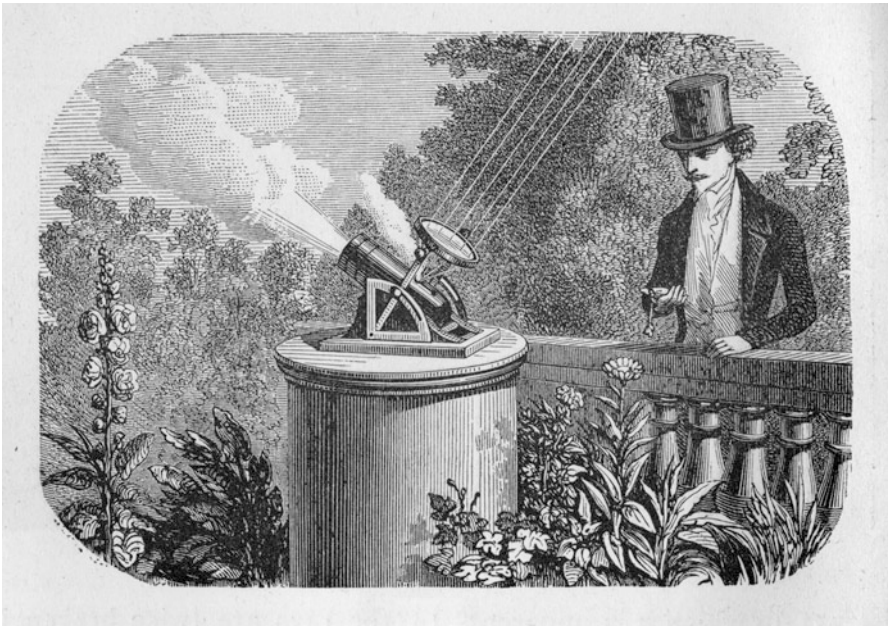


Fig. 8.25 Man setting his watch by the noon blast of a cannon sundial. Adolphe Ganot, *Natural Philosophy for General Readers and Young Persons* (New York, 1872)

Fig. 8.26 A slave to the clock. Heliochronometer by Johann Jakob Sauter II, Stockholm, 1802–1811, which mechanically corrects for the equation of time. Adler Planetarium, Chicago, M-302

Fig. 8.27 Authority of mean time. Heliochronometer by Pilkington and Gibbs, Preston, Lancashire, England, as sold by Ross Ltd., London, circa 1910, along with a promotional booklet. Adler Planetarium, Chicago, W-276

The tables were turned with the advent of the railroads, which required a standard time zone along the tracks, and telegraph lines that could send time signals from precise clocks set up in astronomical observatories to central locations. Mean time won out. Sundials became slaves to clocks. Although Pilkington & Gibbs, England, used pastoral images to advertise its heliochronometers in 1910 (Fig. 8.27), they were often to be found at railroad stations. The firm's slogan—"Tis Always Morning Somewhere"—and model names such as "Empire" underscored belief in the British Empire as an "empire on which the sun never sets." (Schechner 2017, AP: W-276).

Conclusion

The foregoing examples make it clear that sundials tell us much more than the time of day. They tell us about the times in which their makers and users lived. As material culture, the instruments embody social norms, cultural differences, and aspirations related to race, empire, labor, education, fashion, religion, and politics. They show how astronomy's authority was culturally bounded when it came to knowing what the hour was from the sun, moon, and stars.

References

Adler Planetarium, Chicago, Illinois. Adler Collections. Last modified 2016. http://www.adlerplanetarium.org/collections/ [Abbreviated as AP in the text]

The Carthusian Order, The Carthusian Order. Last modified 2016. http://www.chartreux.org/en/

T. Merton, *The Silent Life* (Farrar, Strauss, and Cudahy, New York, 1957)

S.J. Schechner, The material culture of astronomy in daily life: sundials, science, and social change. J. Hist. Astron. **32**, 189–222 (2001a)

S.J. Schechner, in *The Discovery of Time*, ed. by S. McCready, The time of day: marking the sun's passing (MQ Publications, London, 2001b), pp. 120–139

S.J. Schechner, The adventures of Captain John Smith, Pocahontas, and a sundial. The Compendium **14**(1), 19–24 (2007)

S.J. Schechner, The adventures of Captain John Smith among the mathematical practitioners: cosmology, mathematics, and power at the time of Jamestown. Rittenhouse **21**(2), 126–144 (2009a)

S.J. Schechner, Counterfeits, copycats, and knockoffs in the branding and selling of scientific instruments. Paper presented at the XXIII International Congress of History of Science and Technology, Budapest, July, 2009b

S.J. Schechner, *Time and Time Again: How Science and Culture Shape the Past, Present, and Future* (Collection of Historical Scientific Instruments, Harvard University, Cambridge, 2014)

S.J. Schechner, in *How Instruments Have Changed Hands*, ed. by A.D. Morrison-Low, S.J. Schechner, P. Brenni, European pocket sundials for colonial use in American territories, vol. 5 (Scientific Instruments and Collections, Leiden, Brill, 2016), pp. 119–170

S.J. Schechner, *Sundials and Time Finding Instruments of the Adler Planetarium*, vol 1 (Adler Planetarium, Chicago, 2017)

W. Shakespeare, *Shake-Speares Sonnets, Never Before Imprinted* (London, 1609)

L.T. Ulrich, I. Gaskell, S.J. Schechner, S.A. Carter, *Tangible Things: Making History Through Objects* (Oxford University Press, New York, 2015)

Chapter 9
The History of Time

Dennis McCarthy

Abstract Time and timekeeping are subjects of long-standing discussions in the fields of science and philosophy. Timekeeping requires some regularly repeatable natural phenomenon, e.g., a year or a day along with established conventions regarding (1) the interval between repetitions, (2) when the series of repetitions is recognized to begin, (3) the names for successive repetitions and their subdivisions, and (4) the means by which this information can be distributed. Beginning with early man, repetitive astronomical phenomena provided obvious answers to these concerns, the regular appearance of the Sun and Moon being the most obvious. Evolving civilization made it necessary to develop increasingly sophisticated concepts regarding time as well as increasingly sophisticated measurements of the passage of time. Particularly notable steps in the evolution of concepts of time involve the thoughts of Aristotle, Newton, and our understanding of the principles of special and general relativity. Particularly notable events in timekeeping range from sundials, water clocks, mechanical clocks, to atomic clocks, all of these devices being related to increasingly sophisticated astronomical observations.

Keywords Time • Timekeeping requirements

D. McCarthy (✉)
U. S. Naval Observatory, Contractor, International Bureau for Weights and Measures, Sevres, France
e-mail: dennis_mccart57@hotmail.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_9

Chapter 10
"When?" It's a Basic Question That We Ask All the Time

Harlan Stenn

Abstract When did "it" happen? When will we do "it"? At face value, this is pretty simple and easy stuff. For many social interactions, getting the time right to within a few minutes is just fine. There are many billing applications that need better accuracy, as they charge in six-second increments. How important might accurate time be in some medical and health-care situations? In the final nine-seconds before the total power failure that hit the US northeast and Canadian east coast in 2003, about 18,000 events happened in the power distribution grid. The myriad of clocks used to record these events were pretty much unsynchronized and nowhere near accurate. So the logged information was basically useless.

The "When" question is answered with a "time stamp." A time stamp might be a note in your calendar (dinner at 6 pm) or the "Date" entry in an email message. We've all been using a very simple and basic time stamp to record the time since long before we had computers. We have the knowledge and experience to engineer, implement, and deploy a much better time stamp. High-quality time stamps improve the usefulness of data. Better data means things like advances in health care, the ability to make better decisions, better efficiency, reduced liability, and better profitability and financial management.

This talk will explore some of the problems with existing time stamps and the consequences of these problems. To learn about NTF's solution to these problems, its General Timestamp API and Library, and how it can help, visit http://www.nwtime.org.

Keywords Timestamps • General Timestamp API • GTSAPI • Network Time Foundation • SCADA

H. Stenn (✉)
Network Time Foundation, Talent, Oregon, US
e-mail: stenn@nwtime.org; http://www.nwtime.org

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_10

Chapter 11
Inter-site Alignments of Prehistoric Shrines in Chaco Canyon to the Major Lunar Standstill

Anna Sofaer, Robert Weiner, and William Stone

Abstract In and near Chaco Canyon, New Mexico—the center of monumental ceremonial architecture of the Ancestral Puebloan culture—ancient peoples appear to have intentionally interrelated numerous small masonry structures on alignments to the major standstill moon. The structures include low-walled C-shaped, circular, and cairn configurations located on prominent positions near the tops of three mesas that form the south side of Chaco Canyon and mesas located beyond the canyon, with inter-site alignments spanning 5–15 km. Deposits of turquoise and other offerings at these small sites, and their similarity with later Puebloan features, suggest their use as shrines. Geographic information system analysis of the spatial distribution of these sites—with precise geodetic coordinates determined by the National Oceanic and Atmospheric Administration's National Geodetic Survey—shows clustering of their interrelationships along azimuths to the rising and setting moon at its major standstill. Previous extensive investigation by the Solstice Project, with geodetic support by the National Geodetic Survey, documented the Chacoans' commemoration of the lunar standstill cycle at the Sun Dagger petroglyph site on Fajada Butte and in the wall alignments and inter-building relationships of numerous Chaco Great Houses. Other research documented the relationship of the Chacoan Great House of Chimney Rock, Colorado, to the major lunar standstill. Our findings of the inter-shrine-site alignments to the major standstill moon provide significant evidence for a hitherto undocumented small scale of lunar astronomical expression of the Chaco culture. The placement of

A. Sofaer
Solstice Project, Santa Fe, NM 87501, USA

R. Weiner
Solstice Project, Santa Fe, NM 87501, USA

Haffenreffer Museum of Anthropology, Brown University, Providence, RI 02912, USA

W. Stone (✉)
National Oceanic and Atmospheric Administration's National Geodetic Survey,
Santa Fe 87502, NM, USA
e-mail: william.stone@noaa.gov

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_11

these shrines also appears to possibly have marked a correspondence between the topographic trajectory of Chaco Canyon and the alignment to the moon at its major standstill, suggesting a specific effort to integrate Chaco's land formations with celestial patterns. These preliminary findings are part of a study in progress of cosmographic expressions throughout the Chacoan cultural region.

Keywords Chaco Canyon • Archaeoastronomy • Major lunar standstill

Introduction

In this paper, we present evidence that numerous masonry structures located on elevated topographic positions in and near Chaco Canyon—the monumental center of the Ancestral Puebloan world from ca. AD 850 to 1150—were intentionally interrelated on alignments to the major lunar standstill. Previous work by the Solstice Project and others has documented expression of knowledge of the lunar standstill cycle by the ancient Chaco culture at the Sun Dagger petroglyph site (Sofaer et al. 1982; Sofaer and Sinclair 1983) and in the alignments of Chacoan Great Houses (Sofaer 2007; Malville 2004). The 12 sites involved in this study include low-walled, C-shaped, circular, and cairn configurations located on prominent positions near the tops of three mesas that form the south side of Chaco Canyon and mesas located up to 15 km beyond the canyon. The presence of ritual deposits such as turquoise at these sites and formal analogy with features of the descendant Puebloan cultures suggest a primary purpose as shrines. Here, we report on the results of a collaborative effort by the Solstice Project and the National Oceanic and Atmospheric Administration's National Geodetic Survey (NOAA/NGS) to (1) determine the precise geodetic coordinates of these shrine-sites and (2) use geographic information system (GIS) technology to plot this data in real-world space and, in accompaniment with cluster analysis, to investigate possible astronomical inter-site relationships. Our findings suggest that the ancient Chacoans intentionally positioned 12 shrine-sites considered in this study on five inter-site alignments to the rising and setting positions of the moon at its major standstills. In addition, the placement of these shrine-sites may have also marked a correspondence between the topographical alignment of Chaco Canyon itself and the bearing between the rising of the southern major standstill moon and the setting of the northern major standstill moon.

The Chacoans appear to have placed value on the commemoration of the repeating cycles of the sun and moon—encoding their rise and set azimuths of mid-positions and extremes in architectural features, light markings, and inter-site alignments. The resulting spatial order and hierophany of light and shadow events are expressions of the "science of time" which—while complexly interrelated—are in marked contrast to ancient astronomical recordings in codices, tablets, or other written documentation.

The Lunar Standstill Cycle

The moon's standstill cycle is longer (18.6 years) and more complex than the sun's annual cycle, but its rhythms and patterns also can be observed in its shifting positions on the horizon, as well as in its relationship to the sun. In its excursions each month, the moon shifts from rising roughly in the northeast to rising roughly in the southeast and from setting roughly in the northwest to setting roughly in the southwest, but a closer look reveals that the envelope of these excursions expands and contracts through the 18.6-year standstill cycle. In the year of the major standstill, this envelope is at its maximum width, and at the latitude of Chaco, the moon rises and sets approximately 7° north and south of the positions of the rising and setting solstice suns. These positions are the farthest to the north and south that the moon ever reaches. In the year of the minor standstill, 9–10 years later, the envelope is at its minimum width, and the moon rises and sets approximately the same angular distance within the envelope of the rising and setting solstice suns. The progression of the sun and the moon in their cycles can also be quite accurately observed in their changing heights at meridian passage and in the accompanying shifts in shadow patterns. A number of factors, such as parallax and atmospheric refraction, can shift and broaden the range of azimuth where the risings and settings of the solstice suns and the standstill moons appear on the horizon. In addition, judgments in determining a solar or lunar event introduce uncertainties. These judgments involve determining which portion of the object to sight on (e.g., lower limb, center of body, upper limb) and what time to sight it in its rising or setting, as well as identifying the exact time of a solstice or a standstill. In this study, we evaluate inter-site shrine alignments for their possible relationships to the moon's major standstill at both the southern major standstill, when it rises in the southeast at −53.5° and sets in the southwest at 53.5°, and the northern major standstill, when it rises in the northeast at 53.5° and sets in the northwest at −53.5°. These standstill rises/set azimuths of ±53.5° were computed by standard formulation described in Wood (1978); they were checked against tables computed by Anthony Aveni and provided by one of the anonymous reviewers of this paper.

In this paper, we follow the convention established in Sofaer (2007) in which we use astronomical azimuth values that show the reflective (i.e., close to, but not exactly, symmetrical) relationships of opposing rising and setting azimuths of the solar and lunar events. We use 53.5° (east of 0°) for the moon rising at northern major standstill and for the moon setting at southern major standstill and −53.5° (west of 0°) for the moon rising at southern major standstill and setting at northern major standstill. It should be noted that these calculated azimuth values, which take into account standard adjustments for parallax (+0.95°) and atmospheric refraction at the horizon (−0.55°), have been averaged between the computed azimuths of the moon's rising and setting at its northern and southern extremes, because the relationships of the rising and setting azimuths of the solstice suns and the minor and major standstill moons are slightly asymmetric in their relationships to 0° and 180° (note: this averaging has the effect of removing the impact on azimuth of the

choice of moon's tangency [upper limb, lower limb, center of moon] for rise/set events, thereby eliminating the uncertainty accompanying this subjective choice). We do not attempt to specify the direction from which the inter-shrine alignments were conceived (i.e., viewed from the rising or setting direction), and we do not know which directional orientation of alignments would have been more or less significant for the Chacoans. Therefore, it seems especially appropriate to maintain the averaged value, as established in Sofaer (2007).

It is of further note that we do not claim visibility between the shrine-sites as a necessary condition to establish the astronomical significance of their inter-site relationships. Most of the shrines reported here to be interrelated on lunar alignments are not intervisible on these alignments, and in many instances, they do not have visibility to the standstill moon at its rise or set major standstill azimuths on the true/sensible horizon. This does not discount the significance or intentionality of their lunar inter-site alignments. Certain inter-building relationships extend the lunar bearings of central buildings on alignments to outlying buildings across topography that blocks visibility (Sofaer 2007). The ability to maintain an astronomical orientation between non-intervisible points is especially evident in the Chacoan Great North Road, which was aligned on a northern bearing over a distance of 50 km and terminates at a badlands canyon not visible from the roads' origin in Chaco Canyon nor from most positions along its length (Sofaer et al. 1989). Preliminary proto-surveying studies and field experiments by the Solstice Project and others demonstrate various methods (e.g., "wiggling in") by which inter-site astronomical alignments between non-intervisible sites could be established without modern technology within the accuracies of the inter-shrine alignments reported here (Williams et al. 1989; Lekson 2015:153 fn6–154 f. 6,7).

Background Research on the Archaeology and Archaeoastronomy of the Chaco Culture

Chaco Canyon, the center of the Ancestral Puebloan world from ca. AD 850 to 1150, has received considerable attention for its unprecedented size and complexity in the ancient American Southwest (Fig. 11.1). In an attempt to understand Chaco's sociopolitical organization and social hierarchy, scholars have intensively studied the canyon's Great Houses and other monumental architecture which demonstrate massive labor investments despite housing very few inhabitants (Lekson 2006). Currently, Chaco Canyon is widely understood to have been a center of ceremonial and ritual practice for the Ancestral Puebloan people (Kantner 2004; Van Dyke 2007a) rather than playing a redistributive, economic role as posited by previous scholarship (Judge 1989, but see Lekson 2015). Chaco Canyon held influence over a 100,000 sq.km region throughout which 150–200 Great Houses emulated the architectural principles of the central canyon (Marshall et al. 1979; Powers et al.

Fig. 11.1 Southeast aerial view of Chaco Canyon and shrine-site mesa tops at *right*, Great House Pueblo Alto in foreground, and Fajada Butte (location of Sun Dagger site on its summit) in far background. © Adriel Heisey

1983). Vast amounts of pottery—as well as more than 220,000 timbers from 70 to 90 km distance (Toll 2006)—were imported into this inhospitable canyon in an apparently one-way trade relationship in which Chaco seems not to have exported economic goods, but rather religious experience (Renfrew 2001). Various attempts have been made to understand the ideology and practices through which Chaco derived its power and influence, including the positioning of architecture to facilitate visibility toward striking landscape features such as mountain peaks and buttes in a commemoration of "sacred geography" (Van Dyke 2007a); sensory ritual practices that utilized Mesoamerican exotic imports such as cylinder vessels

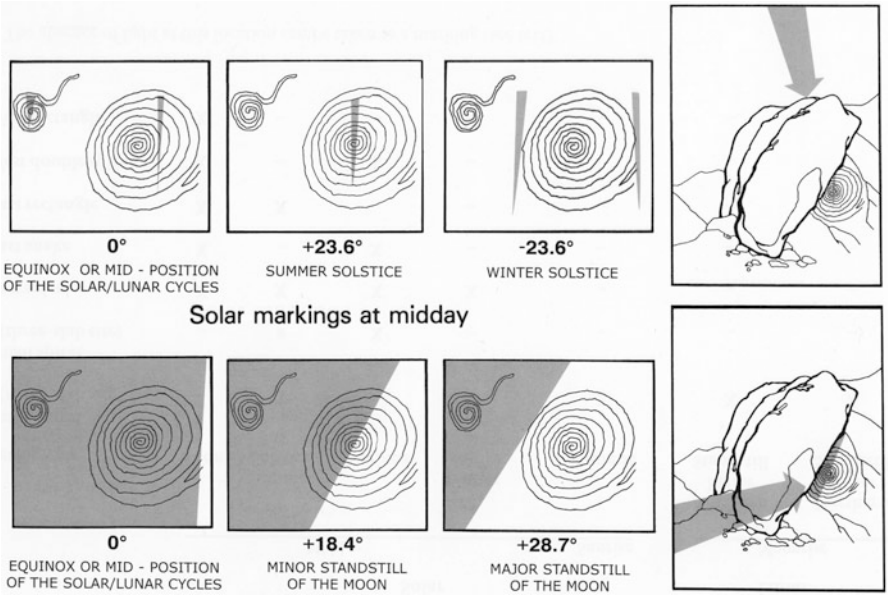


Fig. 11.2 Diagram of shadow and light markings on the Sun Dagger petroglyph at key points in the solar and lunar cycles; sun and moon declinations in degrees. © Solstice Project

containing cacao, scarlet macaws, copper bells, and elaborate shell jewelry[1] (Weiner 2015); and complex astronomical knowledge communicated through light and shadow hierophanies on petroglyphs (Sofaer and Sinclair 1983) and the alignment of buildings (Sofaer 2008).

Previous extensive investigations by the Solstice Project meticulously documented the Chacoans' knowledge of solar and lunar astronomy. Sofaer and colleagues have demonstrated that at the Sun Dagger petroglyph site on Fajada Butte in Chaco Canyon, three upright slabs of sandstone channel sunlight to produce a series of iconic light markings on a spiral petroglyph on the days of summer solstice, winter solstice, and the equinoxes (Fig. 11.2; Sofaer et al. 1979). This same petroglyph also records the lunar standstill cycle with a shadow cast on the spiral by the rising moon at its northern major and minor standstills (Fig. 11.3; Sofaer et al. 1982).

Subsequent work by the Solstice Project, with survey support of NOAA/NGS, revealed that the major walls of numerous Great Houses in Chaco Canyon and the wider region are aligned to cardinal, solstitial, and lunar standstill azimuths (Sofaer 2007).[2]

[1]We note that the astronomical knowledge demonstrated by the Chaco culture may have been another form of "exotica" the Chacoans acquired in their dynamic trade/ideological relationship with Mesoamerican cultures (Sofaer 2008:xviii; Tuwaletstiwa 2015:xv; Weiner 2015).

[2]We acknowledge that our previous work on Chaco building alignments has been critiqued (e.g., Malville 2008; Munro and Malville 2011) and these concerns will be addressed in future work. For a recent third-party publication in support of the Solstice Project's previous findings, we recommend Pauketat (2016).

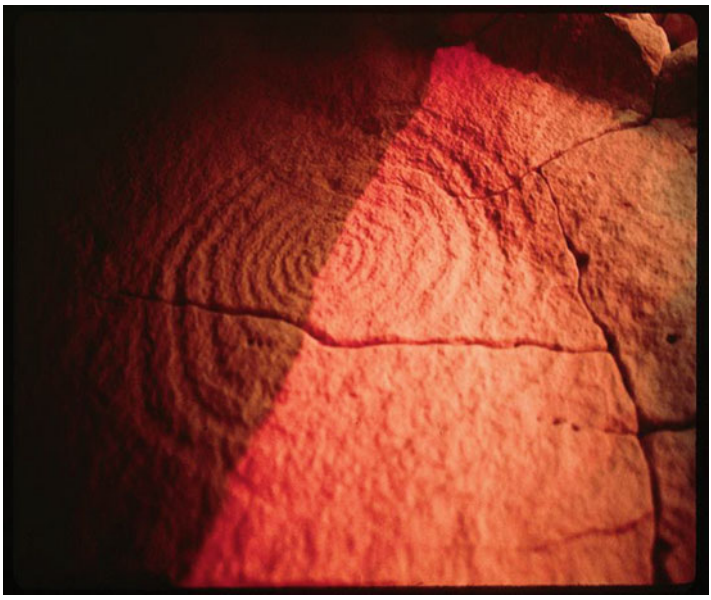


Fig. 11.3 Simulated shadow of moonrise at northern minor lunar standstill at the Sun Dagger site. © Solstice Project

Other evidence for commemoration of lunar astronomy among the Chacoans is documented at the outlying Great House of Chimney Rock in the foothills of the Rocky Mountains near Pagosa Springs, Colorado −135 km from Chaco Canyon. Malville (2004:138–139) demonstrated that the Chimney Rock Great House is positioned to have visibility to two large rock pillars that perfectly frame the rising moon at its northern major standstill, and tree ring cut dates confirm construction of the Great House in years of major standstills (AD 1018, AD 1076, and AD 1093)[3] (see also Lekson 2015:195–196). These previous documentations of the Chaco culture's numerous commemorations of the lunar standstill cycle initiated our interest in the inter-site alignments of 12 shrine-sites on elevated positions of the mesas in and near Chaco Canyon (see Sofaer 2007:250n24).

Crescents, Cairns, and Circles in Chaco Canyon

While the majority of research has attempted to understand the Chaco culture and their expressions of astronomical knowledge through monumental architecture, small archaeological features such as mesa-top masonry crescents, circles, and

[3]It is of interest to note that an additional cut date from Chimney Rock dates to AD 1011, the year of the minor lunar standstill (Lekson 2015:196).

cairns have been less studied (but see Hayes and Windes 1975; Van Dyke 2007a; Van Dyke et al. 2016). Building on several earlier surveys, the Solstice Project comprehensively documented masonry crescentic, circular, and cairn sites throughout Chaco Canyon and the surrounding landscape in the 1980s with archaeologist Michael Marshall.

Site Forms

The stone crescents, circles, and cairns of our study are located on elevated landforms such as mesa cliff edges, butte tops, and mesa summits in and near Chaco Canyon. We wish to emphasize that the sites included in this study are solely those found on prominent elevated positions and it is their placement at high elevations that unites these various forms of crescents, circles, and cairns. We follow Van Dyke et al. (2016) in categorizing the various Chacoan feature types of crescent, circle, and cairn features as shrines. Crescentic site sizes range from 3.5 m to 18.0 m across the long axis (Figs. 11.4 and 11.5; Marshall and Sofaer 1988:254). The majority of the crescentic structures open toward the east and southeast. Their walls are normally low (50 cm to 1 m tall), and it is apparent these partial enclosures were open and unroofed. While ceramics are sparse at crescentic sites, the wares that are found at them suggest temporal affiliation with

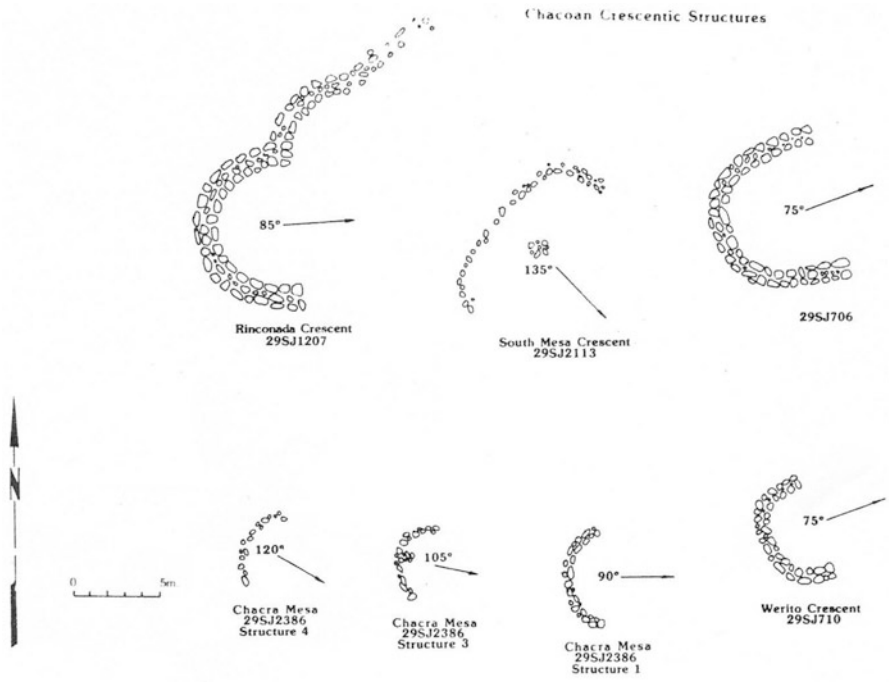


Fig. 11.4 Maps of the seven crescentic sites discussed in this paper (Marshall and Sofaer 1988)

Fig. 11.5 Crescentic shrine-site LA 40710 and view to the west. Photo by William Stone, NOAA's National Geodetic Survey

the late Pueblo I through early Pueblo III periods[4]—contemporary with the pinnacle of Chaco Canyon's regional influence (Marshall and Sofaer 1988:264). Marshall and Sofaer (1988:256) note: "the celestial exposure and cycloramic perspective from these elevated locations is clearly a principal factor to be considered in any attempt to decipher the function and significance of crescentic structures." C-shaped and circular low-walled masonry features have served as shrines in later Puebloan contexts (Fowles 2009; Harrington 1916; Ortiz 1969; Snead 2008:81–112), suggesting a possible ceremonial association with these Chacoan features.

Also included in this study are a site of interlinked, small-walled circular structures (the Poco Site) and three sets of cairns[5]—cylindrical stacks of

[4]The most common types include Red Mesa Black-on-white, Escavada Black-on-white, Gallup Black-on-white, Chaco-McElmo Black-on-white, and Cibola plain, banded, and corrugated indented (Marshall and Sofaer 1988:264).

[5]Two other small cairn sites on Chacra Mesa (29SJ184 & 29MC463), located approximately 3 km southeast of our study area, have also been brought to our attention (Ruth Van Dyke, personal communication Jan 7, 2016). It is of interest that preliminary examination, using locational information provided by Van Dyke, shows that both appear to be on the inter-site same alignments to the lunar standstill presented in this paper.

Fig. 11.6 One of 12 cairns at shrine-site LA 41088 on cliff edge of West Mesa with view to the west. © Anna Sofaer

rocks—also located on elevated locations. The cairns at LA 41088 have been noted for their impressive size (up to 1.5 m in diameter and 1.5 m in height) and Bonito-style masonry known in Chaco buildings (Fig. 11.6; Marshall and Sofaer 1988:238–244). Since the mesa tops of Chaco Canyon have been comprehensively surveyed over five decades of archaeological study, it is probable that the sites included in our study are the full set of such elevated sites in and near Chaco Canyon.

Site Functions

The presence of ritual deposits such as turquoise and projectile points at these sites (Marshall and Sofaer 1988), their careful masonry construction, and analogy with later Puebloan features suggest their having a primary purpose as shrines. For example, investigations at the crescent located at 29SJ423 [LA 40423] revealed a buried "stone bowl . . . covered by a large stone slab" that held 146 turquoise beads, found frequently as ritual deposits in burials and sealed in kiva niches at other sites throughout Chaco Canyon[6] (Fig. 11.7; Hayes and Windes 1975:149). However, two

[6]LA 40423 also is located next to a Basketmaker III site that predates the Bonito-era occupation of Chaco Canyon and might be understood as attempting to express a link with real or imagined ancestors (Van Dyke 2007a:142).

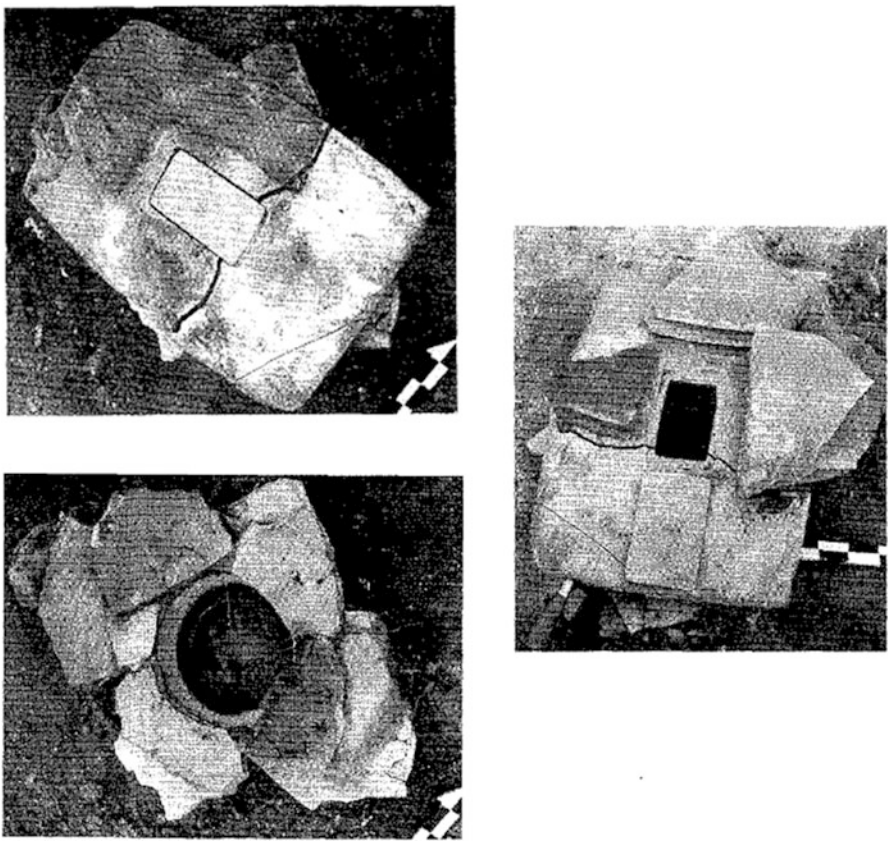


Fig. 11.7 Turquoise-filled repository at LA 40423 (Mathien 2005: Fig. 5.24)

other hypotheses have been offered to explain the primary function and location of these shrine-like sites: that they served to establish visibility both between themselves and with Chacoan buildings (Hayes and Windes 1975) and that they served as signaling stations for outlying Great Houses in the Chaco region (Van Dyke et al. 2016). It is certainly possible that some of the features of our study—especially those located in the greater landscape beyond Chaco Canyon—may have been involved in intervisibility networks. Most provide dramatic views of the larger landscape of the San Juan Basin and, in some cases, Great Houses or other shrines (Marshall and Sofaer 1988; Van Dyke 2007a). However, several points suggest that most of the features of our study likely did not have signaling as their primary purpose. No evidence for fires or ash pits (i.e., to produce smoke signals or nighttime light) has been found at these sites (Marshall and Sofaer 1988:255). While it is proposed that the Chacoans may have used selenite (a reflective mineral that is found frequently in the San Juan Basin) to signal from shrines, the potential times for signaling would be curtailed to only the specific times of day when the sun

was properly positioned. We note also that the study area of 15 km is of such a limited distance and the crescentic structures so closely spaced that messages within the canyon and nearby landscape could be more easily carried by a runner than through smoke signals or signaling with selenite. Despite these observations, the shrine features in and near Chaco Canyon certainly could have played more than one role in Chacoan society. The primary purpose of our analysis is not so much to consider the various proposed uses of these features, however, as to test whether there is a significant underlying pattern to their spatial distribution.

As noted earlier, this analysis includes all the masonry crescents, circles, and cairns known at the time of this study that are positioned on elevated locations within and near Chaco Canyon—i.e., within the 15 km of our study area. Only two of the sites of our study are located on the north mesa of Chaco Canyon: most are located on the mesas forming the south side of Chaco Canyon. There are other known circular stone features, but these are located on significantly lower elevations and exclusively on the mesas forming the north side of Chaco Canyon. Furthermore, the northern mesa stone circles constitute a different class of architectural features, consisting of circles of single stones rather than constructed low walls in distinct contrast to the sites of this study (Windes 1978). Additionally, herraduras—U-shaped/crescent masonry features frequently found in association with Chaco roads and especially in the wider Chaco region (Nials et al. 1987)—are considered a separate feature class and not included in this study, as the features considered here are not associated with roads, include more and different types of votive offerings, and are located in dramatically high topographic positions on the mesas of the central canyon.

The features included in this study consist of seven low-walled, crescentic masonry structures on elevated mesa tops/edges (Peñasco Crescent [LA 40423], South Crest Structure [LA 40706], Werito Crescent [LA 40710], Rinconada Crescent [LA 41207], South Mesa Crescent [LA 42113], Chacra Mesa Crescent [LA 42386], and LA 42859), a set of two stone cairns (Gateway Cairns [LA 51167]), a series of 13 cairns (Poco Cairns, [LA 51489]), one circular rock alignment located on top of an isolated mesa approximately 15 km from Chaco Canyon (Pretty Rock [LA 37676]), one circular rock alignment with 12 nearby stone cairns located on the edge of a steep mesa (LA 41088), and a site consisting of three inter-locking circular rock alignments (the Poco Site [LA 41010]) (Table 11.1).

Methods

The shrine-sites considered in this study were documented by the Solstice Project during archaeological fieldwork in 1984 and 1985 (Marshall and Sofaer 1988). All sites were drawn to scale, their probable cultural-temporal affiliation determined,

Table 11.1 Shrine-sites in and near Chaco Canyon of this study

Site number	Name	Type
LA 40423	Peñasco Crescent	Crescent
LA 40706	South Crest Structure	Crescent
LA 40710	Werito Crescent	Crescent
LA 41207	Rinconada Crescent	Crescent
LA 42113	South Mesa Crescent	Crescent
LA 42386	Chacra Mesa Crescent	Crescent
LA 42859	n/a	Crescent
LA 51167	Gateway Cairns	Two cairns
LA 51489	Poco Cairns/Chaco North Cairn complex	13 cairns
LA 37676	Pretty Rock	Circle
LA 41088	Chaco West Cairn complex	Circle and 12 cairns
LA 41010	Poco Site	Inter-locking circles

and their topographic positions and visibility to other landscape features and archaeological sites were described, along with any artifacts present. In subsequent years, precise geodetic coordinates were obtained for 11 of the 12 sites studied here during three fieldwork sessions in 2007, 2009, and 2014. Site coordinates were determined by GPS observations collected with survey-grade equipment and processed through the NOAA/NGS Online Positioning User Service (http://geodesy.noaa.gov/OPUS/). Three other shrines (Pretty Rock, LA 41088, and the Gateway Cairns) could not be GPS occupied at the time of this study, so their latitude and longitude were obtained using Google Earth. A digital elevation model (DEM) was constructed for the study area by combining enhanced 10 m grid LiDAR rasters freely available from the New Mexico Resource Geographic Information System website (http://rgis.unm.edu). The DEM and site locations were then imported into Esri's ArcMap 10.2 for analysis.

Next, a Python script, based on the formulation of the NOAA/NGS geodetic inverse tool (http://geodesy.noaa.gov/TOOLS/Inv_Fwd/Inv_Fwd.html), was written to calculate the azimuths between every possible combination of shrine-site pairs and to help determine if a predominant cluster of inter-site azimuths appeared and whether such a cluster might correspond with any significant solar or lunar alignment. The data derived from these calculations were plotted into a histogram using Microsoft Excel for visual analysis.

Line features were then drawn in ArcMap to connect those series of shrines which, based on the inter-site azimuth calculations and histogram analysis, appeared to be positioned on inter-site relationships to the major lunar standstill. After first using the NOAA/NGS geodetic inverse tool to compute inter-site azimuths, it was determined that Google Earth's measure tool provided equally accurate calculations of inter-site azimuths; thus, Google Earth was used to determine the azimuths presented in this paper.

Results

The table of all possible azimuths of alignments between shrine-site pairs, when plotted in a histogram, presents a clear clustering of azimuths near −53.5°, corresponding with the alignment between the rising of the southern major standstill moon and the setting of the northern major standstill moon (Appendix; Fig. 11.8). In other words, the most common inter-site azimuth is clustered at −53.5°, with 14 pairs (see pairs bolded in Appendix). Five azimuths of inter-site alignments close to ±53.5° are presented in Table 11.2 and are shown plotted onto a GIS map in Fig. 11.9. Four of these five azimuths fall within 1.1° of the major standstill SE-NW moonrise/set (i.e., −53.5°), with the other falling within 1.6° of the major standstill NE-SW moonrise/set (i.e., 53.5°)—a level of precision that would be difficult to obtain with a modern compass, although consistent with certain of the alignments found in Chacoan building orientations (Sofaer 2007: Fig. 9.6).

The inter-shrine alignments to the major standstill as proposed here also appear to be associated with another architectural feature type—isolated great kivas. Great kivas are monumental, semisubterranean circular chambers widely thought to have served a ceremonial and integrative function in ancient Chacoan society, measuring, in the case of Casa Rinconada (LA 814), 19.5 m in diameter (Fig. 11.11; Van Dyke 2007b). While many great kivas occur in the courtyards of Great Houses, four are located in isolated positions unassociated with Great Houses. There are only two isolated great kivas in or close to the central Chaco Canyon complex (Van Dyke 2007b:112–116), and both are located on the lunar inter-shrine alignments found in this study. Isolated great kiva 1253 (LA 41253) falls on the lunar standstill

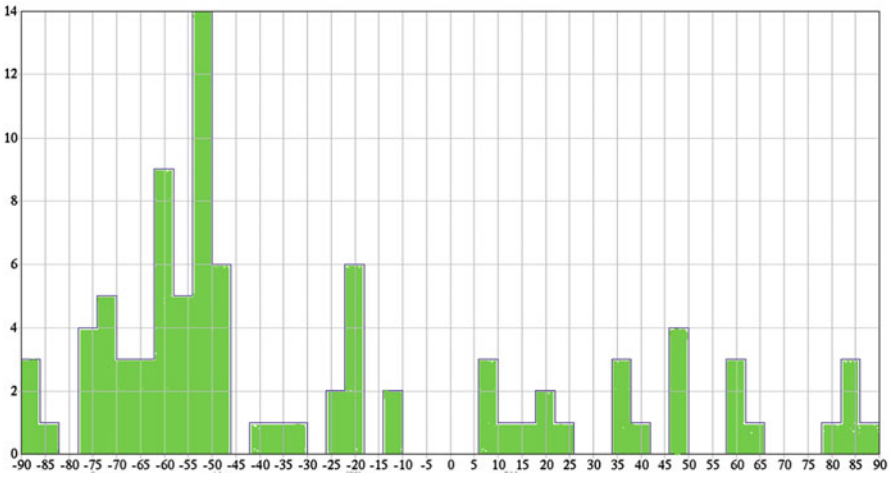


Fig. 11.8 Histogram of inter-site alignments of all possible pairings of shrine-sites/cairns and clustering at the major lunar standstill −53.5° azimuth. Plot by Demetrios Matsakis (US Naval Observatory)

Table 11.2 Measured geodetic azimuths of inter-site alignments of crescentic, circle, and cairn sites in Chaco Canyon as compared with calculated ideal lunar azimuths

Sites involved in inter-site alignment	Calculated ideal lunar azimuth	Measured geodetic azimuth of inter-site alignments	Difference between calculated ideal lunar azimuth and measured inter-site bearings
A. Pretty Rock (LA 37676) to LA 40423 to Casa Rinconada (LA 814) to LA 42859	−53.5°	−53.6°	0.1°
B. Gateway Cairns (LA 51167) to LA 41088 to LA 42113 to LA 40710	−53.5°	−54.6°	1.1°
C. LA 40423 to LA 40706	−53.5°	−52.6°	0.9°
D. LA 42386 to isolated great kiva 1253 (LA 41253) to LA 41207	−53.5°	−53.5°	0°
E. Isolated great kiva casa Rinconada (LA 814) to the Poco Site (LA 41010), passing through the center of the 13 Poco Cairns (LA 51489)	53.5°	55.1°	1.6°

Sites are listed in order from northwest to southeast (except alignment E, the NE-SW major standstill alignment, lists sites from southwest to northeast). Alignments A–E are also labeled on map (Fig. 11.9)

alignment from the Chacra Mesa Crescents to the Rinconada Crescent at a bearing of −53.5°, and the isolated great kiva Casa Rinconada is positioned on the lunar standstill alignment between the Poco Cairns and Poco Site at a bearing of 55.1° (alignment E, Fig. 11.9). It is also of interest that the Peñasco Blanco Great House, which itself is aligned to the rising of the southern major standstill moon, is located on an inter-site alignment with Casa Rinconada—an alignment which parallels most of the inter-shrine alignments discussed in this paper[7] (Sofaer 2007:250n25).

In summary, our preliminary findings using cluster analysis support that crescentic, circular, and cairn sites in and near Chaco Canyon were intentionally positioned on inter-site alignments to the major lunar standstills.

Discussion and Conclusion

Cultures across ancient America incorporated cosmological principles in their architecture and settlement patterns (e.g., Ashmore and Sabloff 2002; Aveni 1980; Broda 2015; Pauketat 2013; Snead and Preucel 1999). Recent studies suggest that the ancient inhabitants of Cahokia (Pauketat 2013; Pauketat et al. 2017), the

[7]In addition, the Great House Una Vida, which itself is aligned to the setting of the southern major standstill moon, is also positioned on an alignment with isolated great kiva 1253 at a bearing of 55.3°, paralleling the inter-site alignment of the Poco Site to Casa Rinconada.

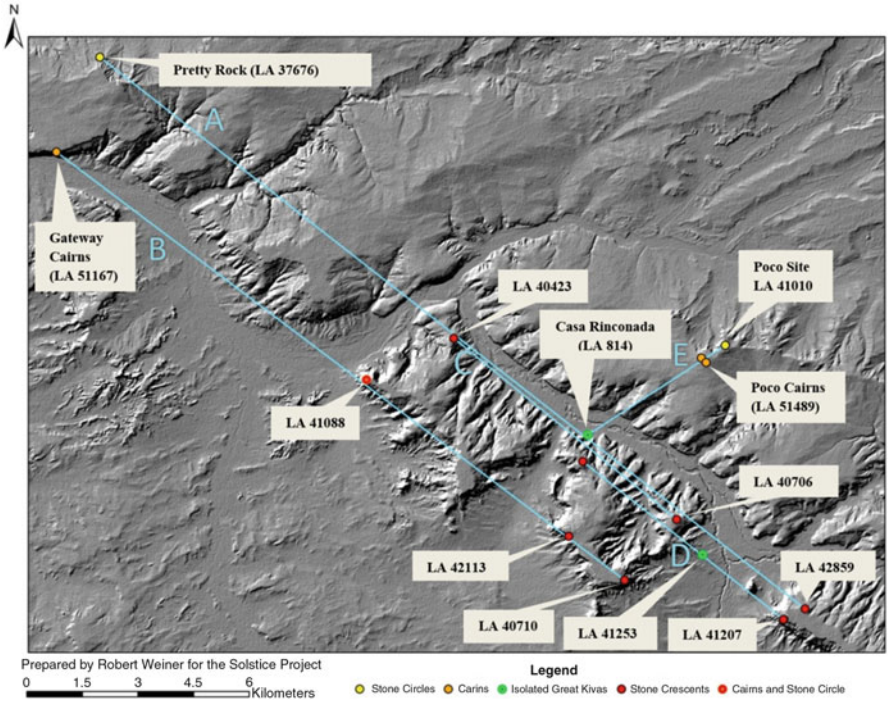


Fig. 11.9 LiDAR terrain map of Chaco Canyon, with survey-grade GPS positions of crescentic, circle, and cairn sites and proposed inter-site alignments to the major lunar standstill. The letters labeling alignment lines correspond with those provided in Table 11.2. Popular names are given for sites that are referred to by their popular name throughout the text

Hopewell culture (Hively and Horn 2013), and the ancient Maya people of Palenque (Mendez et al. 2005) and the Maya Lowlands (Šprajc 2016) oriented their monumental works to the lunar as well as the solar cycles. Relating physical structures to these celestial cycles appears to have been especially important to the Chaco culture, who built multistoried Great Houses in a pattern of alignments to the cardinal directions and lunar standstill azimuths (Sofaer 2007). A particular interest in lunar astronomy and the standstills is also evidenced among the ancient Chacoans in light and shadow markings at the Sun Dagger petroglyph site (Sofaer et al. 1982; Sofaer and Sinclair 1983), wall and inter-building alignments of Great Houses (Sofaer 2007), and the positioning of the Chimney Rock Great House in view of two rock spires that frame the major standstill moonrise (Malville 2004). It has been proposed that these astronomical alignments and commemorations were also essential to the timing of ceremonial events and pilgrimage into Chaco Canyon and that control over this regionally shared calendar was one source of Chaco's power (Judge and Malville 2004; Sofaer 2007). The multiplicity and diversity of these commemorations of lunar astronomy by the Chaco culture offer further support for the intentionality of the lunar inter-shrine-site alignments proposed here.

Fig. 11.10 Shrine-site 1088 and northwest view of setting northern major standstill full moon over Chaco River, January 4, 2007. © William Stone

Certain key numbers associated with the archaeological features involved in the inter-site alignments also appear to underscore their lunar significance. LA 41088 consists of 12 cairns in association with its stone circle (Fig. 11.10), and the Poco Cairns site consists of 13 cairns—numbers which correspond closely with the lunar months (12.3) in the solar year. The proposed alignment to the rising of the southern major standstill moon between the Poco Site and Casa Rinconada may also be supported by evidence from the internal features of Casa Rinconada. It has 28 niches—the number of days in the lunar month—evenly positioned around its circular walls (Fig. 11.11).[8] Great kiva niches often held votive deposits of turquoise, shell, or other ideologically charged materials, suggesting their ritual importance in ancient Chacoan society (Van Dyke 2007a:125).

[8]Astronomer Carl Sagan drew a correspondence between Casa Rinconada's 28 niches and the number of days in the lunar cycle (Sagan 2013:43).

Fig. 11.11 Casa Rinconada, an isolated great kiva, located in central Chaco Canyon. Note its position on two major lunar standstill inter-site alignments. It is of interest that its niches of ritual offerings number 28. © Anna Sofaer

The symbolic significance of the Chacoans' expressions of solar and lunar astronomy is suggested by parallels in descendant Puebloan cultures.[9] Puebloan people have for centuries commemorated and observed the sun and moon for ritual and calendric purposes (Ellis 1975; McCluskey 1977), often at small stone sites similar to those found in the Chaco region (Stevenson 1904:355). Such sites are typically located on elevated positions and at significant distances from the pueblos and sometimes in the form of small crescentic and circular stone structures.

Phillip Tuwaletstiwa, geodetic scientist and member of the Hopi tribe, offers insight into the significance of the Chacoans' astronomical alignments of its central complex from the perspective of Puebloan traditions. He sees the alignments as "a

[9]While there are many parallels of the Chacoan culture with the traditions of historic and contemporary Puebloan cultures, there are also distinct differences. For instance, the Chacoans' multistoried Great Houses with up to 400–800 rooms and engineered roads of 9 m width are not present in the historic Pueblos. Similarly, there are no known historic Puebloan alignments to the 18.6 year standstill cycle of the moon. It has been frequently stated that Pueblo people deliberately left Chaco and the knowledge and hierarchy manifested there for a choice to live in a more "reciprocal relationship with Mother Earth" (Ortiz 1992:72; interviews in Sofaer 1999). For these reasons, we do not find it necessary or appropriate to refer to ethnographically documented Puebloan practices for validation of the findings of astronomical expression by the Chaco culture.

way to transfer the orderly nature of the cosmos down onto what seems to be chaos that exists here ... to then integrate both heaven and earth—and this would be ... viewed typically in Pueblo culture as a center place" (Sofaer 1999). The late anthropologist Alfonso Ortiz of Ohkay Owingeh noted the importance of the moon in Puebloan cosmology (predicting the presence of lunar markings at the Sun Dagger site) because of its complementary relationship with the sun (Alfonso Ortiz, personal communication 1978). He also guided us to the significance of land formation and elevated sites in Puebloan cosmology as essential insight into the cosmographic implications of the findings of this study (Ortiz 1969; Alfonso Ortiz, personal communication 1986).

It is of special interest that the lunar inter-shrine alignments reported here parallel the land formation of Chaco Canyon itself, revealing the canyon's approximate topographic alignment to the major standstill moon. This distinctive correspondence of the canyon's trajectory with the moon's cycle may have been an important factor in the Chacoans' decision to commemorate the lunar standstill cycle on the tops of the mesas forming the south side of the canyon. Such deliberate acts to integrate land formations with the moon's cycle would have been a profound unification of land and sky.

Numerous instances have been documented of the Chaco people's labor-intensive investments to incorporate land formations in their astronomical alignments as well as in their road alignments. The 50 km Great North Road of the Chaco culture appears to have been built with the purpose of both commemorating the direction north and articulating with special topographic features by terminating at the steepest descent into the badlands of Kutz Canyon (Sofaer et al. 1989). The 57 km South Road ends at towering Hosta Butte, a singularly prominent feature on Chaco Canyon's southern horizon (Sofaer et al. 1989; Marshall 1997). Chimney Rock Great House is set on a narrow, sharp precipice to incorporate the view of the major standstill moon rising between two massive rock pillars (Malville 2004). The Sun Dagger site and other petroglyph light markings are positioned remotely near the top of the 135-m-high Fajada Butte (Sofaer et al. 1979; Sofaer and Sinclair 1983). In this context, it would not be surprising that in addition the Chacoans noted and marked the canyon's topographic alignment to the moon's cycle with the placement and interrelationships of shrines on high sites. We suggest that the Chacoans' inspiration to so extensively commemorate the lunar standstill at Chaco Canyon may have been in part inspired by the canyon's apparently close alignment to the moon's major standstill azimuths. This paper is part of a larger Solstice Project study currently in progress to explore these concepts of Chacoan cosmography.

The Chaco culture's expression of the "science of time" appears to commemorate astronomical cycles in ways distinctive from written documentations by other ancient societies. First, the Chacoans marked repeating cycles in azimuth positions rather than in tabulated periodicities or the documentation of historic astronomic events (e.g., the appearance of comets, meteors). Second, they translated temporal celestial cycles into a spatial order on the landscape. Third, they did so through the

establishment of a geometric plan in the layout of Chaco Canyon's central complex to the rising and setting of the key positions in the solar and lunar cycles.

In this paper, we have presented evidence that crescentic, circular, and cairn sites in and near Chaco Canyon were intentionally interrelated on alignments to the major lunar standstill. A cluster analysis of all possible inter-site alignments suggests that the spatial distribution of these sites appears to be organized in relation to the setting and rising positions of the moon at major standstill. Our findings of the inter-shrine alignments to the major standstill moon provide significant evidence for a hitherto undocumented small scale of lunar astronomical expression of the Chaco culture, in parallel with its large-scale architectural alignments. Further support is provided by the lunar numbers built into some of the shrine-sites. In future work, we propose to evaluate the significance of the inter-site alignments of crescentic, circular, and cairn shrines in the larger context of the Chacoans' cosmography and ritual landscapes (e.g., Marshall 1997; Sofaer 2007; Sofaer et al. 1989; Stein and Lekson 1992; Van Dyke 2007a). Such research has the potential to further illuminate the many ways this ancient culture integrated astronomy, the natural landscape, and architecture in cosmographic expressions across a vast region of the American Southwest.

Acknowledgments We owe a special debt to NOAA's National Geodetic Survey for supporting the precise survey of the shrine-sites of this study, following its critical support of the Solstice Project's earlier survey of the major Chacoan Great Houses. Archaeologist Michael Marshall conducted his usual thorough and insightful surveys of the sites, describing them in the context of his extensive knowledge of such sites throughout the Chaco region. Demetrios Matsakis (US Naval Observatory) generously produced the histogram that shows the clustering of the inter-shrine alignments and calculated moon rise and set azimuth data, which are fundamental and crucial to our analysis. We also greatly appreciate the support of the National Park Service and the special attention that archaeologist Dabney Ford gave to the process of this study. Adriel Heisey conducted aerial photography of the alignments and sites of this study, in one instance flying in his ultralight plane in the freezing post-blizzard conditions that preceded the mid-winter full moon of 2007 at the major standstill moon. We also thank geodesist Phillip Tuwaletstiwa for initiating our work with the National Geodetic Survey and for encouraging this further study of the significance of the moon to the Chaco culture. The late Alfonso Ortiz, anthropologist and member of Ohkay Owingeh Pueblo, guided our early attention to the special importance of the moon in Puebloan cosmology and of the symbolic significance of elevated small sites. The late Rina Swentzell of Santa Clara Pueblo, on observing our preliminary findings, noted that the lunar inter-shrine alignments—and especially their correspondence with the canyon land formation—were "brilliant." We thank Micah Lau for his assistance writing the Python script to calculate all possible inter-site azimuths. We were also helped and encouraged by archaeologists Ruth Van Dyke and Richard Friedman. Two anonymous reviewers helped us to make important clarification to our text and the computation of moon rise and set azimuths and encouraged us to explore further the cultural implications suggested by these findings. All three authors of this paper contributed significantly to its creation.

Appendix

A list of all inter-site azimuths (reckoned from north at 0° and expressed as ± degree bearings east/west of 0° that acknowledge the reflective [i.e., close to, but not exactly, symmetrical] relationships of the rising and setting azimuths of northern and southern major lunar standstill positions) between the 12 sites included in this study, with the cluster within ±1.5° of the major standstill azimuth (−53.5°) bolded.

6.94	46.42	−18.47	−51.16	**−54.61**	−65.06
8.85	60.53	−19.03	−51.20	**−54.75**	−66.55
13.83	61.23	−19.13	−51.68	−55.37	−66.89
17.22	61.84	−21.11	**−52.13**	−56.73	−67.38
18.68	62.77	−23.51	**−52.15**	−58.06	−70.41
21.37	79.33	−24.06	**−52.54**	−58.84	−70.45
34.76	82.78	−31.78	**−52.89**	−58.86	−70.94
36.45	84.53	−36.87	**−53.19**	−59.96	−73.70
37.48	85.37	−41.29	**−53.35**	−60.20	−73.95
47.01	87.62	−46.00	**−53.42**	−60.98	−74.57
48.98	−8.77	−46.70	**−53.43**	−61.18	−75.53
49.40	−13.60	−46.98	**−53.54**	−61.68	−77.74
89.72	−13.70	−47.93	**−53.77**	−61.71	−77.77
22.67	−18.37	−48.39	**−53.87**	−64.01	−82.66
38.66	−18.46	−48.51	**−54.04**	−64.89	−86.26
					−87.23

References

W. Ashmore, J.A. Sabloff, Spatial orders in Maya civic plans. Lat. Am. Antiq. **13**(2), 201–215 (2002)

A.F. Aveni, *Skywatchers of Ancient Mexico* (University of Texas Press, Austin, 1980)

J. Broda, Political expansion and the creation of ritual landscapes: a comparative study of Inca and Aztec cosmovision. Camb. Archaeol. J. **25**(1), 219–238 (2015)

F.H. Ellis, A thousand years of the Pueblo sun-moon-star calendar, in *Archaeoastronomy in Pre-Columbian America*, ed. By A.F. Aveni (University of Texas Press, Austin, 1975), pp. 59–88

S.M. Fowles, The enshrined Pueblo: villagescape and cosmos in the northern Rio Grande. Am. Antiq. **74**(3), 448–466 (2009)

J.P. Harrington, in 29th Annual Report of the Bureau of American Ethnology, The Ethnogeography of the Tewa Indians (Bureau of American Ethnology, Washington, 1916), pp. 29–618

A.C. Hayes, T.C. Windes, in Papers in Honor of Florence Hawley Ellis, pp. 143–56. Papers of the Archaeological Society of New Mexico 2, ed. by T.R. Frisbie, An Anasazi shrine in Chaco Canyon (Archaeological Society of New Mexico, Santa Fe, 1975)

R. Hively, R. Horn, A new and extended case for lunar (and solar) astronomy at the Newark earthworks. Midcont. J. Archaeol. **38**(1), 83–118 (2013)
W.J. Judge, Chaco Canyon—San Juan basin, in *Dynamics of Southwest Prehistory*, ed. By L.S. Cordell, G.J. Gumerman (Smithsonian Institute Press, Washington, 1989), pp. 209–261
W.J. Judge, J.M. Malville, Calendrical knowledge and ritual power, in *Chimney Rock: the Ultimate Outlier*, ed. By J.M. Malville (Lexington Books, Lanham 2004), pp. 151–162
J.W. Kantner, *Ancient Puebloan Southwest* (Cambridge University Press, Cambridge, 2004)
S. H. Lekson (ed.), *The Archaeology of Chaco Canyon: an Eleventh-Century Pueblo Regional Center* (School for Advanced Research Press, Santa Fe, 2006)
S.H. Lekson, *The Chaco Meridian: One Thousand Years of Political and Religious Power in the Ancient Southwest*, 2nd edn. (Rowman & Littlefield, Lanham, 2015)
J.M. Malville, Ceremony and astronomy at Chimney Rock, in *Chimney Rock: the Ultimate Outlier*, ed. By J.M. Malville (Lexington Books, Lanham, 2004), pp. 131–150
J.M. Malville, *A Guide to Prehistoric Astronomy in the Southwest* (Johnson Books, Boulder, 2008)
M.P. Marshall, The Chacoan roads: a cosmological interpretation, in *Anasazi Architecture and American Design*, ed. By B.H. Morrow, V.B. Price (University of New Mexico Press, Albuquerque, 1997), pp. 62–74
M.P. Marshall, A. Sofaer, Solstice Project Investigations in the Chaco District 1984 and 1985: The Technical Report. Manuscript on file, Laboratory of Anthropology, Santa Fe, New Mexico, (1988)
M.P. Marshall, J.R. Stein, R.W. Loose, J. Novotny, *Anasazi Communities of the San Juan Basin* (Historic Preservation Bureau, Santa Fe, 1979)
F.J. Mathien, in *Culture and Ecology of Chaco Canyon and the San Juan Basin*. Publications in Archaeology 18H, Chaco Canyon Studies (National Park Service, Santa Fe, 2005)
S.C. McCluskey, The astronomy of the Hopi Indians. J. Hist. Astron. **8**, 174–195 (1977)
A. Mendez, E.L. Barnhart, C. Powell, C. Karasik, Astronomical observations from the temple of the sun. Archaeoastronomy **19**, 44–73 (2005)
A.M. Munro, J.M. Malville, Ancestors and the sun: astronomy, architecture and culture at Chaco canyon. Proc. Int. Astron. Union **7**(S278), 255–264 (2011)
F.L. Nials, J.R. Stein, J.R. Roney, *Chacoan Roads in the Southern Periphery: Results of Phase II of the BLM Chaco Roads Project* (Bureau of Land Management, Albuquerque, 1987)
A. Ortiz, *The Tewa World: Space, Time, Being, and Becoming in a Pueblo Society* (University of Chicago Press, Chicago, 1969)
S. Ortiz, What we see: a perspective on chaco canyon and pueblo ancestry, in *Chaco Canyon: a Center and Its World*, ed By M. Peck (Museum of New Mexico Press, Albuquerque, 1992), pp. 65–72
T.R. Pauketat, *An Archaeology of the Cosmos: Rethinking Agency and Religion in Ancient America* (Routledge, London/New York, 2013)
T.R. Pauketat, A ray of theoretical sunshine. J. Skyscape Archaeol. **2**(2), 251–254 (2016)
T.R. Pauketat, S. Alt, J. Kruchten, The emerald acropolis: elevating the moon and water in the rise of Cahokia. Antiquity **99**(355), 207–222 (2017)
R.P. Powers, W.B. Gillespie, S.H. Lekson, The outlier survey: a regional view of settlement in the San Juan Basin. Reports of the Chaco Center No. 3. National Park Service, Albuquerque, (1983)
C. Renfrew, Production and consumption in a sacred economy: the material correlates of high devotional expression at Chaco Canyon. Am. Antiq. **66**(1), 14–25 (2001)
C. Sagan, *Cosmos* (Ballantine Books, New York, 2013)
J.E. Snead, *Ancestral Landscapes of the Pueblo World* (The University of Arizona Press, Tucson, 2008)
J.E. Snead, R.W. Preucel, The ideology of settlement: Ancestral Keres landscapes in the Northern Rio Grande, in *Archaeologies of Landscape: Contemporary Perspectives*, ed. By W. Ashmore, A.B. Knapp (Basil Blackwell, Oxford, 1999), pp. 169–197
A. Sofaer (dir), *The Mystery of Chaco Canyon*, (Bullfrog Films, Oley, 1999)

A. Sofaer, The primary architecture of the Chacoan culture: a cosmological expression, in *The Architecture of Chaco Canyon*, ed. By S.H. Lekson (The University of Utah Press, Salt Lake City, 2007), pp. 225–254.

A. Sofaer, Introduction, *in Chaco Astronomy: An Ancient American Cosmology* (Ocean Tree Books, Santa Fe, NM, 2008), pp.xiii–xx.

A. Sofaer, R.M. Sinclair, Astronomical Markings at Three Sites on Fajada Butte, Chaco Canyon, New Mexico, in *Astronomy and Ceremony in the Prehistoric Southwest*, ed. By J. Carlson, W.J. Judge (University of New Mexico Press, Albuquerque, 1983), pp. 43–49. Maxwell Museum of Anthropology, Anthropological Papers No. 2

A. Sofaer, V. Zinser, R.M. Sinclair, A unique solar marking construct: an archaeoastronomical site in New Mexico marks the solstices and equinoxes. Science **206**(4416), 283–291 (1979)

A. Sofaer, R.M. Sinclair, L.E. Doggett, Lunar markings on Fajada Butte, Chaco Canyon, in *Archaeoastronomy in the New World*, ed. By A.F. Aveni (Cambridge University Press, Cambridge, 1982), pp. 169–181.

A. Sofaer, M.P. Marshall, R.M. Sinclair, The Great North Road: a cosmographic expression of the Chaco culture of New Mexico, in *World Archaeoastronomy*, ed. By A.F. Aveni (Cambridge University Press, Cambridge, 1989), pp. 365–376

I. Šprajc, I.: Lunar alignments in Mesoamerican architecture. Anthropological Notebooks **22**(3), 61–85 (2016)

J.R. Stein, S.H. Lekson, Anasazi Ritual Landscapes, in *Anasazi Regional Organization and the Chaco System*, ed. By D. Doyle (University of New Mexico Press, Albuquerque, 1992), pp. 87–100. Maxwell Museum of Anthropology Anthropological Papers 5

M.C. Stevenson, In 23rd Annual Report of the Bureau of American Ethnology, The Zuñi Indians. (Smithsonian Institution, Washington, DC, 1904), pp. 9–157

H.W. Toll, Organization of production, in *The Archaeology of Chaco Canyon: an Eleventh-Century Regional Center*, ed. By S.H. Lekson (School for Advanced Research Press, Santa Fe 2006), pp. 117–151

P. Tuwaletstiwa, Foreword in three movements ii, in *The Chaco Meridian: One Thousand Years of Political and Religious Power in the Ancient Southwest*, 2nd edn. ed. By S.H. Lekson (Rowman & Littlefield, Lanham, 2015), p. xiii-xixv

R.M. Van Dyke, *The Chaco Experience: Landscape and Ideology at the Center Place* (School of American Research Press, Santa Fe, 2007a)

R.M. Van Dyke, Great kivas in time, space, and society, in *The Architecture of Chaco Canyon*, ed. By S.H. Lekson (The University of Utah Press, Salt Lake City, 2007b), pp. 93–126

R.M. Van Dyke, R.K. Bocinsky, T. Robinson, T.C. Windes, Great houses, shrines, and high places: intervisibility in the Chacoan world. Am. Antiq. **81**(2), 205–230 (2016)

R.S. Weiner, A sensory approach to exotica, ritual practice, and cosmology at Chaco canyon. Kiva **81**(3–4), 220–246 (2015)

R. Williams, F.D. Uzes, R. Sinclair, A. Sofaer, Surveying methods the Anasazi may have used to determine, record, and reproduce a direction. Solstice Project Unpublished Report (1989)

T.C. Windes, in Reports of the Chaco Center No. 5, Stone Circles of Chaco Canyon, Northwestern New Mexico, (National Park Service, Department of the Interior, Albuquerque, 1978)

J.E. Wood, *Sun, Moon and Standing Stones* (Oxford University Press, Oxford, 1978)

Chapter 12
Atomic Time Scales and Their Applications in Astronomy

Felicitas Arias

Abstract The unit of time is defined as a multiple of the period of the hyperfine transition of the atom of cesium 133 and realized at the level of a few parts in 10^{16} by about a dozen cesium fountains maintained in national metrology institutes. International Atomic Time (TAI) takes its accuracy from the primary frequency standards, but preserves its high stability over intervals of 1 month making use of the largest industrial clock ensemble in the world. Seventy-four institutes worldwide spread or disseminate atomic time for different applications; the International Bureau of Weights and Measures (BIPM) collects their data and integrates them in an algorithm which produces TAI.

For practical reasons the time scale which has been recommended as the world's time reference is Coordinated Universal Time (UTC), derived from TAI by applying a procedure defined by the International Telecommunication Union (ITU). The institutes and observatories contributing clock data to the BIPM maintain local representations of UTC, indicated by UTC(k), and provide time for all sorts of applications, including those linked to Earth and space science.

For some applications in astronomy, long-term stability is a requisite and cannot be satisfied by the "quasi-real-time TAI." For these users, the BIPM computes yearly a representation of terrestrial time, TT(BIPM), supported by all the primary frequency standards submitted measurements.

This paper will describe the characteristics of the atomic time scales and present the link between international time metrology and the astronomical science.

Keywords Time scales • International Atomic Time (TAI) • Coordinated Universal Time (UTC) • Terrestrial time (TT) • Time metrology • Coordinate times

For practical reasons the time scale which has been recommended as the world's time reference is Coordinated Universal Time (UTC), derived from TAI by apply-

F. Arias (✉)
BIPM, Pavillon de Breteuil, Sèvres Cedex, France
e-mail: farias@bipm.org

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_12

ing a procedure defined by the International Telecommunication Union (ITU). The institutes and observatories contributing clock data to the BIPM maintain local representations of UTC, indicated by UTC(k), and provide time for all sorts of applications, including those linked to Earth and space science.

Chapter 13
Relativistic Time at the US Naval Observatory

Matsakis Demetrios

Abstract Since the nineteenth century, the US Naval Observatory (USNO) has contributed to the science of relativity and also benefited from it. Albert Michelson received USNO support and employment before going on to undertake the Michelson-Morley experiment. USNO first flew atomic clocks around the world as a test of relativity. USNO staff have written articles and a textbook on relativity, and a pedagogical explanation of why curved space-time causes an apple to fall is provided. Two recent tests of Einstein's equivalence principle (EEP) are described, and a section in memory of Dr. Gernot Winkler is included at the end of this work.

Keywords Relativity • U.S. Naval Observatory • Einstein Equivalence Principle • Dr. Gernot Winkler • Apple Falling From Tree

Introduction

The USNO's mission, as indicated on its web pages (http://www.usno.navy.mil/USNO/about-us/the-usno-mission), is the following:

1. Determine the positions and motions of celestial bodies, motions of the Earth, and precise time.
2. Provide astronomical and timing data required by the Navy and other components of the Department of Defense for navigation, precise positioning, and command, control, and communications.
3. Make these data available to other government agencies and to the general public.
4. Conduct relevant research, and perform such other functions as may be directed by higher authority.

The USNO's current mission is similar to earlier mission statements (Dick 2003) and supplemented by acts of law, such as the Naval Appropriations Act of 1849, which called for the publication of the *Nautical Almanac*, and the America

M. Demetrios (✉)
U.S. Naval Observatory, 3450 Massachusetts Avenue, NW, Washington, DC 20392, USA
e-mail: demetrios.matsakis@navy.mil; dnmyiasou@yahoo.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_13

COMPETES Act of 2007, which assigns coordinated responsibility for the definition of Coordinated Universal Time (UTC) to the Secretaries of Navy and Commerce. We will now describe several instances of USNO's involvement in the development and testing of relativity.

Albert Michelson

Perhaps the earliest involvement of the USNO began in 1878, when Simon Newcomb funded future Nobel laureate Albert Michelson (then a young ensign) to work on the speed of light (Dick 2003). This led to a year's stay at the USNO, during which Michelson and Newcomb jointly conducted experiments with fixed mirrors at the USNO, at the Washington Monument, and at Ft. Myers in Virginia. Michelson later teamed with Edward Morley to make the surprising 1887 discovery that the Earth is stationary with respect to the aether. The constancy of the speed of light and the assumption that all observers are equivalent (the relativity of simultaneity) are the cornerstone of Einstein's special relativity theory (Einstein 1905), in that these two assumptions are sufficient to derive the Lorentz-FitzGerald contraction, time dilation, and the famous formula $E = mc^2$. Figure 13.1 shows an inspirational portrait of Michelson that is on permanent display in the USNO Superintendent's office. Since allowing daylight to enter the laboratory would lower the data's signal-to-noise ratio, one may speculate that the picture reflects an acknowledgment of the importance of calibration and alignment.

Relativity for Educating Both the Scientists and the Public

The USNO has contributed to education and technical training of relativity in many ways, and this is driven by the need to incorporate relativity in celestial mechanics, geodesy, astrometry, and timekeeping. Scientists who are interested in correctly applying the theory to problems involving geodesy can consult the IERS Conventions (2010), which were first developed at the USNO and for which the USNO Scientific Director remains an editor. George Kaplan showed that the corrections used to reduce optical data were consistent with those used in the radio (Kaplan 1998). Researchers interested in the application of relativity to celestial mechanics may consult a recently published 860-page work on the subject (Kopeikin et al. 2011).

The Hafele-Keating Experiment

One of the key consequences of the special theory is that the relative change in proper time is not affected by issues of simultaneity. The first demonstration of time

Fig. 13.1 Albert Michaelson

dilation was in the increased lifetime of cosmic-ray muons (Rossi and Hall 1941; Frisch and Smith 1963); however, the demonstration that most captured the public's interest was undertaken as a strictly scientific test when Dr. Gernot Winkler was director of the USNO Time Service Department. In collaboration with Dr. Hafele of the University of Colorado (Fig. 13.2), this experiment was to fly a cesium ensemble around the world in an easterly direction, with an expected loss of 40 ns, and again in a westerly direction, with an expected gain of 275 ns. These predictions agreed with the measurements within the 20 ns uncertainties, which were mostly in the predictions (Hafele and Keating 1972). Various versions of this demonstration have since been repeated as demonstrations for the public many times. The BBC produced a documentary in which John Davis transported a clock round trip from the UK's National Physical Laboratory (NPL) to the USA, and most recently, Tom Van Baak's cesium clocks were driven up a mountain to demonstrate gravitational time dilation for the "GENIUS" series by Steven Hawking (http://www.pbs.org/about/blogs/news/pbs-and-national-geographic-channels-international-commission-six-part-series-genius-by-stephen-hawking/ (Episode 1)).

An Illustration of the Principle of Least Action

For pedagogical purposes, we now note that it is possible to provide an equationless explanation for why an apple falls from the tree, given that gravity is considered explainable as an illusion related to the presence of curved space. Without the classical force of gravity, a student might wonder why an apple would "choose to fall." We invoke the slightly misnamed Principle of Least Action, also informally known as the "Law of Cosmic Laziness." In analogy with the classical Principle of

Fig. 13.2 J. Hafele (*left*) and R. Keating

Least Action, this leads to equations that require light and matter to follow the geodesic, which is the straightest possible path in curved space-time. The principle can be reformulated as saying that the proper time (exemplified in the apple's "biological clock") is minimized over the path.

The first step in this simplified explanation is to realize that if there magically appeared a hole just beneath the apple tree, extending all the way through the Earth, the apple would fall through the Earth to the other side and then return to the branch from which it fell. This would be a one-dimensional orbit, as shown in Fig. 13.3.

The question now becomes why would the apple go through this degenerate orbit, rather than simply sit on the tree? Figure 13.4 depicts the two ways the apple could travel, one of them hypothetical. In each case the apple moves from the first event (when it is on the tree and starts to fall) to the second event, in which it is again on the tree. If a uniform density is assumed, the time to reach the Earth's center is about 42 min; more realistic models compute values about 10% less (Klotz 2015).

We now make the assumption that minimizing the action (proper time) over each segment of the apple's path through its fall (a local effect) is equivalent to minimizing the action over the entire path (a global effect); in complex gravitational fields, this would not necessarily be the case. As the distance separation between the start and end points of the two possible paths is zero, the relative change in proper time is not obscured by the issue of defining simultaneity. Therefore, the reason the apple chooses to fall is because it will have aged less with respect to a hypothetical apple that stayed fixed to the tree (and not in free fall). The reason it would have aged less is gravitational time dilation – clocks in a gravitational field go slower. Large slowdowns involving black holes are the substance of movies; in our case the average frequency dilation is of order 10^{-9}. We therefore find that the apple is willing to transverse four Earth radii (2.4×10^4 km) so that after its return about 10^4 s later, its proper time will age less by about 10^{-5} s.

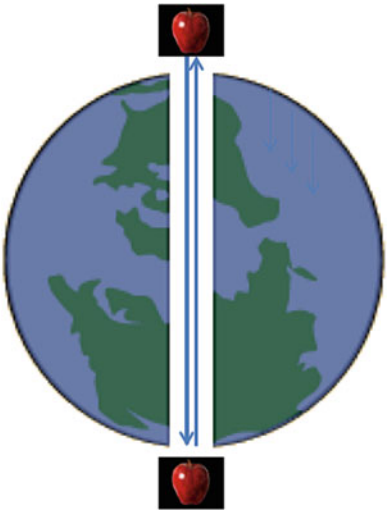


Fig. 13.3 If it didn't hit a hard surface, the apple would travel through the Earth and back

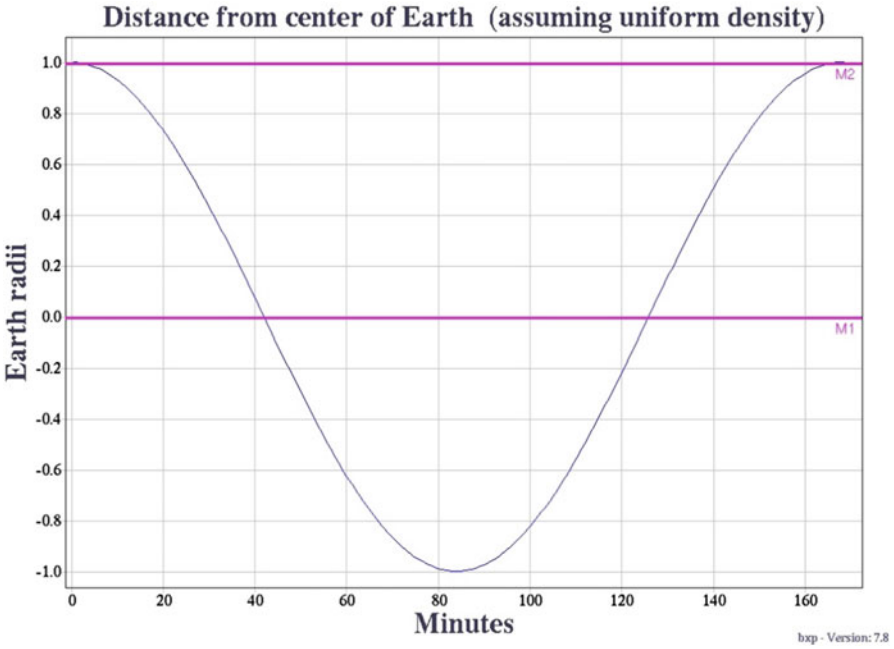


Fig. 13.4 The two possible paths an apple can take between two events at which it is at the tree. The horizontal line at 1.0 Earth radius is the (hypothetical) path in which it remains stationary. The second path takes the apple through a hole in the Earth and back, generating the sinusoid in the figure

USNO Tests of Einstein's Equivalence Principle (EEP)

Einstein's equivalence principle is at the foundation of relativity. It states that an observer cannot, by local measurements, distinguish between a gravitational field and acceleration. It follows that there always exists an accelerating frame for which the laws of special relativity apply locally and in which gravitation should be entirely ignored. This principle has several variants (Will 2014). The strong EEP is that this principle holds even when the energy of gravitation is considered. Lorentz invariance applies to the motion of the observer, and Local Position Invariance (LPI) stipulates that EEP holds anywhere in the universe (and therefore in any gravitational field). The Schiff conjecture is that a violation of any of these variants of EEP implies a violation of the others; this conjecture can be proven if one is willing to assume the conservation of energy (Lightman and Lee 1973). Many theories of quantum gravity predict violations of EEP at some level, and therefore, every test has the potential to constrain these theories.

Search for EEP (LPI) Violations in Annual Solar Potential Variations

Piel et al. (2013) looked for dependence on the solar gravitational potential in frequency differences between the four USNO rubidium fountains, the USNO cesium ensemble, and elements of the USNO hydrogen maser ensemble. Since the solar potential, as experienced by the Earth, varies with a 1-year periodicity, any frequency variations in phase with this potential could be ascribed to an atom-dependent violation of EEP. They found no variations, to a 1-sigma upper limit of 1.3×10^{-6} in the LPI-violating parameter β for hydrogen-cesium and 1.7×10^{-5} for rubidium-hydrogen. Figure 13.5 shows best fits to an annual variation in the frequency differences, scaled for display. Figure 13.6 shows how the USNO data have reduced the upper limits to possible variations in the fine structure constant, the electron/proton mass ratio, and the quantum chromodynamic parameter ratio m_q/Λ_{QCD}.

Searches for EEP (LPI) Violations in Diurnal Solar Potential Differences of the Earth

We now describe a test for EEP violations that utilized satellite-based frequency transfer between widely separated hydrogen masers (Matsakis 2016a), whose frequencies might be expected to have an almost diurnal signature as the Earth's

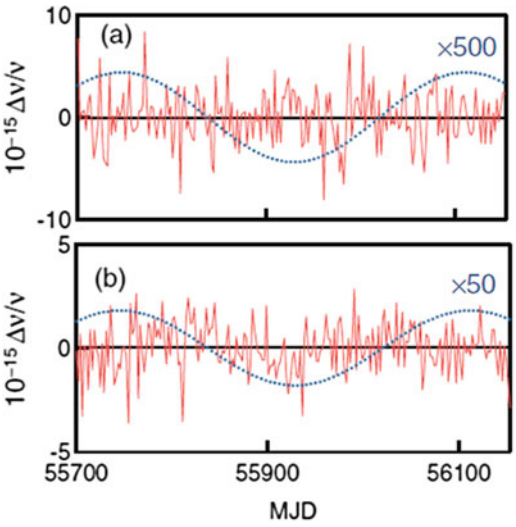


Fig. 13.5 Best fits to annual term in frequency difference between different types of clocks, scaled as shown for display. The *upper plot* is one hydrogen maser minus the cesium ensemble, and the *lower plot* gives one hydrogen maser minus the rubidium fountain average. Figure extracted from Piel et al. (2013)

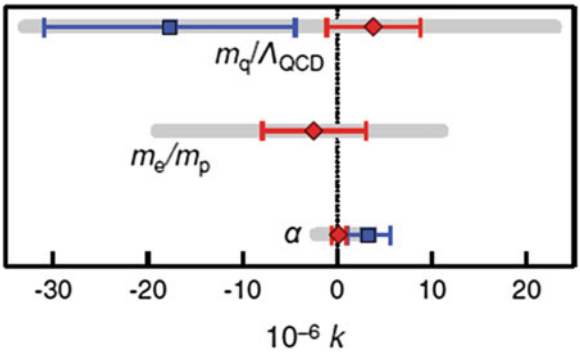


Fig. 13.6 Constraints on the coupling of three ratios of physical constants to the solar gravitational potential. The *gray bands* show the constraints from previous measurements, the *squares* show the constraints imposed using only USNO data, and the *diamonds* show the tighter constraints when all data are combined. Figure extracted from Piel et al. (2013)

rotation carries the clocks deeper into and further out of the solar gravitational potential well. However, it has long been known that there is no observable "noon-midnight" redshift in GPS; this is due to the fact that all clocks are essentially in free fall about the Sun. It was noted by many authors, such as Hoffman (1961) that the large redshift that might be expected in a ground clock and a satellite at noon (roughly four times the 36 ns/day expected between the ground clock and the center of the Earth, ignoring the Earth's gravity) is numerically canceled by the second-order Doppler shift expected from the difference between the satellite and the ground clocks' Keplerian orbits around the Sun. While no one disputes that EEP brings about this instantaneous cancelation, it has been argued in Ashby and Weiss

(2013) that the subsequent motion of the satellite and Earth invalidates this approach because at a later time the cancelation occurs in a different accelerated frame than what is the appropriate frame at the time in question. The authors then show that the nonexistence of a frequency difference can be derived by a more general approach using the relativity of simultaneity.

In Matsakis (2016a), frequency transfer was obtained in two different ways. One was by differencing about 4 years of time transfer data obtained by two-way satellite time transfer (TWSTT, also known as TWSTFT) between the USNO's Washington DC facility and its Alternate Master Clock (AMC), in Colorado Springs, Colorado. In all, there were three USNO antennas and two AMC antennas employed, resulting in six baselines. An independent series of frequency transfer data was obtained using 8 years of monthly Precise Point Positioning (PPP) solutions made publicly available by the International Bureau of Weights and Measures (BIPM). The PPP solutions are in the form of differences with a common reference, IGS time (IGST). The frequency difference between pairs of ground clocks was obtained through subtraction, and this eliminated IGST from the analysis.

For frequency transfer involving TWSTT, a single transponder on a geostationary satellite was employed. Therefore, relativistic effects in the satellite are largely irrelevant. However, PPP data are reduced with clock and orbit models computed by the International GNSS Service (IGS), and the sets of satellites observed by clocks on opposite sides of the Earth have at most only a few satellites in common. This difference was ignored in Hoffman (1961) because the details of how the IGS computes its products are difficult to extract from their average of analysis center products. Since a proper allowance for this effect might be expected to increase the average magnitude of the observed frequency differences, the author concluded that the PPP-derived upper limits may be more stringent than claimed.

For both data sets, frequency transfer differences were edited by baseline (pair of receivers) to remove outliers. The PPP solutions were edited by day; and any baseline that had even one outlier was completed removed for that day. The next step was to compute the expected shift due to the solar potential alone (which would be canceled by the instantaneous second-order Doppler), given the Sun-Earth separation on each day. The expected shift was taken as a parameter to describe the EEP violation. For reference, an orthogonal second parameter was also fitted. It was equal in magnitude to the EEP-violation parameter but offset in time by 6 h. Although both of these parameters are orthogonal to a constant frequency difference, a receiver-pair-dependent constant was removed from each day's data before the fit, and then two parameters were solved for using all receiver pairs for which data were available on that day. For both TWSTT and PPP, the correlations due to common ground clocks were treated using standard pre-whitening techniques. Pre-whitening requires knowledge of the nonwhite behavior, which is not exactly known. Therefore, reductions were made using either the correlations in the "noise" or an idealized model. For each pre-whitening mode, the sets of 1-day parameter fits were edited to remove 3-sigma outliers (about 7% of the solutions). The fitted parameters were then averaged into 10-day bins, and again 3-sigma outliers were

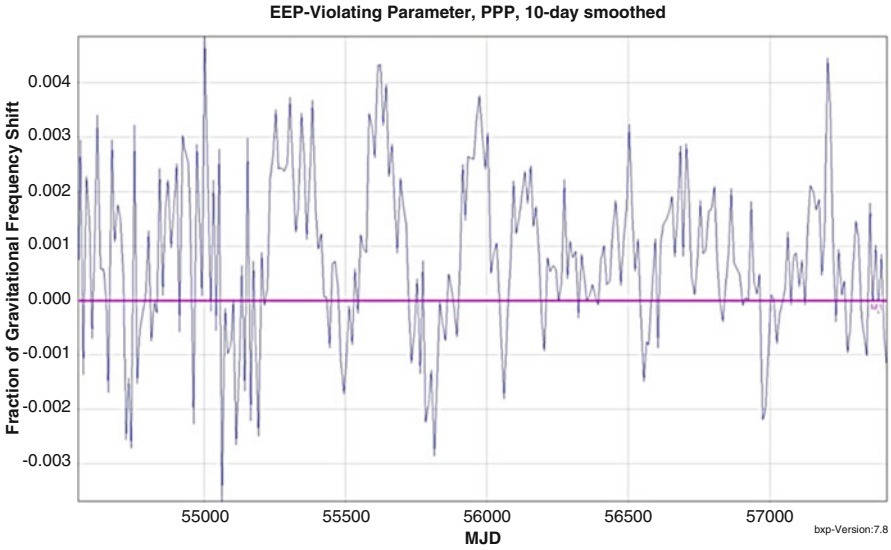


Fig. 13.7 The EEP-violating parameter as a function of time. Note the nonwhite behavior and the improvement in recent data

removed (one or two points per mode). Some results are shown in Figs. 13.7 and 13.8, and all are summarized in Table 13.1. In Table 13.1, the statistical uncertainties are computed using the standard Gaussian formula for the standard deviation of the mean; however, it is obvious that the data are nonwhite. Since the recent PPP data are less nonwhite and quieter (reflecting improvements in the technology), averaging was also done excluding the older data.

Since the uncertainties are merely formal computations, the high "signal-to-noise" ratio of some of the solutions does not imply an EEP violation. The quadrature parameter shows almost as large of a variation, and this would be expected to be zero even if an EEP violation existed. We ascribe this variation to diurnal signatures in both the PPP and the TWSTT data. The diurnal variations in TWSTT data are well known, and while they will likely be strongly reduced through the use of a software correlator (Matsakis 2016b), in this data set, they vary both in amplitude and with regard to time of day (Huang et al. 2016). At those times when they are in phase with the Sun, they would be indistinguishable from an EEP violation. The diurnal variations in PPP are weaker; however, at times they are strong enough to be directly observed. Figure 13.9 provides an example of data between a laboratory in Italy and one in Japan. These are not typical of all the data, but they are typical of times when there were strong diurnal variations. The diurnals shown were in all baselines involving the receiver IENG, which was removed from the analysis. The diurnals associated with other PPP baselines are weaker than shown here and often not directly detectable.

Since diurnals due to unrelated causes would obscure the EEP variation, it was concluded that a reasonable upper limit to the EEP violations would be $<10^{-3}$. This

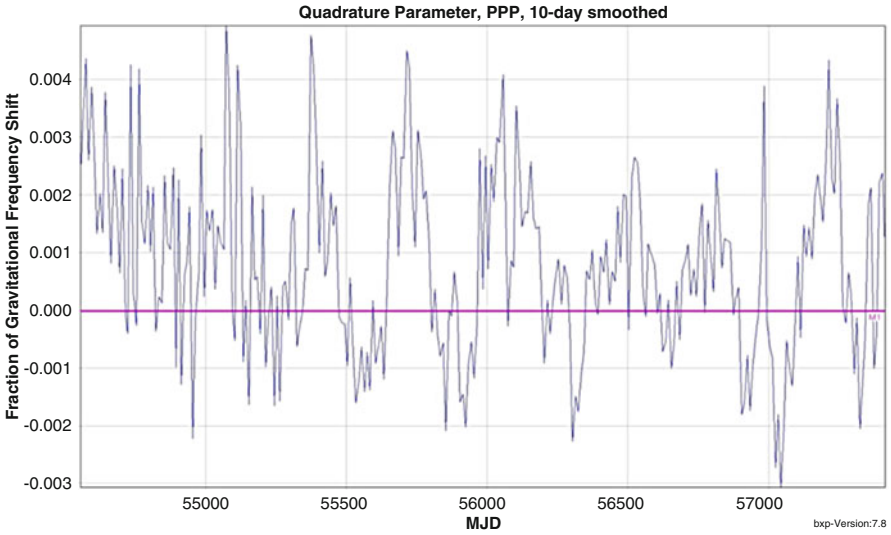


Fig. 13.8 The fitted value of the quadrature parameter, which is orthogonal to the EPP-violating parameter

Table 13.1 Properties of fits to the EEP-violating and quadrature parameters, for different solution and frequency transfer techniques

Technique	# of days	Covariance source	EEP-violating parameter	Statistical uncertainty	Quadrature parameter
TWSTT	1538	Data	-2.2×10^{-4}	$\pm 2.4 \times 10^{-3}$	5.0×10^{-4}
TWSTT	1538	Model	1.4×10^{-3}	$\pm 2.4 \times 10^{-3}$	5.5×10^{-3}
PPP	2865	Data	7.4×10^{-4}	$\pm 1.1 \times 10^{-4}$	4.4×10^{-4}
PPP	2865	Model	7.9×10^{-4}	$\pm 8.9 \times 10^{-5}$	7.8×10^{-4}
PPP	1217	Data	6.8×10^{-4}	$\pm 1.1 \times 10^{-4}$	2.8×10^{-4}
PPP	1217	Model	6.1×10^{-4}	$\pm 9.2 \times 10^{-5}$	4.4×10^{-4}

The last row summarizes the most stringent test and provides the justification for claiming the limit of $<10^{-3}$

would make it the most accurate test to date using widely separated clocks in the solar field (Fig. 13.10).

In the near future, more accurate tests of the EEP in the solar field are expected from the use of optical fibers and optical clocks (Rasel 2015), both of which are very close to achieving operational use. The USNO also has an optical clock development project underway, with a goal of exceeding the frequency precision of its four rubidium clocks by an order of magnitude, perhaps to 10^{-17} s/s. Rockets carrying clocks into orbits passing near the Sun have always had the potential of highly accurate relativity tests; in their case it is simply a matter of when they will be funded.

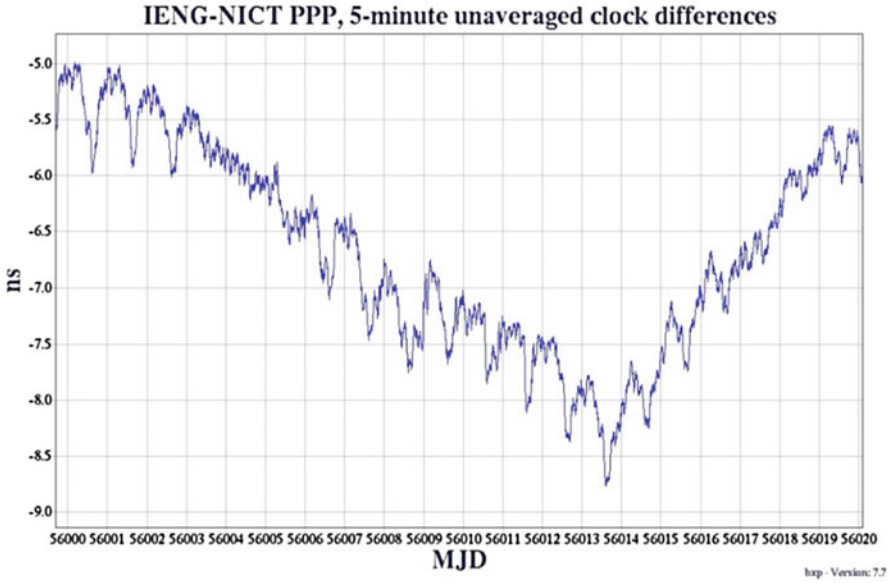


Fig. 13.9 Unedited PPP data between laboratories in Italy (IT) and Japan (NICT). The vertical bars mark the start of each day (UTC = 0). The plot shows a period of strong diurnals. Similar diurnals were seen in all baselines involving IENG at these times, and this site was excluded from the analysis

Conclusion

Relativity, since its inception a fascinating field of fundamental physics, has over the last several decades become just as active in applied physics. The USNO has contributed to this field in the past and continues doing so to this day.

In Memory of Dr. Gernot Winkler

Gernot Winkler served as director of the USNO Time Service Department for 30 years, from 1966 to 1996 (Fig. 13.11). He was brilliant, dynamic, inspirational, and successful. On his second day in office, he made his first improvement to use a cesium clock instead of a quartz clock as the USNO Master Clock. For the next 30 years, he continued to improve not just his department but the timekeeping community in many ways, such as by providing development funding, kick-start funding, and/or operational development of masers, digital 5071 cesium clocks, mercury stored-ion clocks, portable clock trips, TWSTT, and Earth rotation studies through the world's largest photographic zenith tube (PST), connected element interferometry (CEI), and very long baseline interferometry (VLBI). He and Louis Essen independently conceived of the now-controversial concept of leap seconds

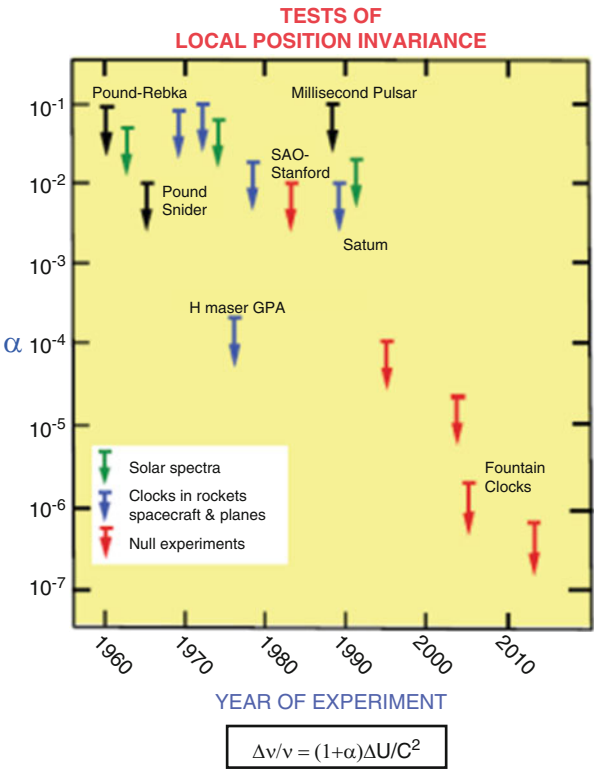


Fig. 13.10 Upper limits to LPI violations, from Will (2014)

Fig. 13.11 Dr. Winkler with an early model cesium clock

Fig. 13.12 Gernot Winkler in retirement

(http://tycho.usno.navy.mil/papers/ts2014/MatsakisLeapSecondComments-URSI-2014.pdf or http://www.gps.gov/cgsic/meetings/2014/), and this is what enabled atomic time, through UTC, to become the world time standard. He successfully implemented programs so that UTC, via the USNO, would be the time reference for LORAN, transit, and GPS. He also contributed significantly to discussions concerning the operational implementation of relativistic corrections in GPS (Fig. 13.12).

Dr. Winkler was always extremely interested in philosophy. In his later years, he used the Internet to publish his ideas in the form of essays on the philosophy behind physics, science, society, and morality. They can be found in http://gmrwinkler.net.

References

N. Ashby, M. Weiss, Why there is no noon-midnight shift in the GPS (2013), http://arXiv [gr-qc]:1307.6525

S.J. Dick, *Sky and Ocean Joined. The U.S. Naval Observatory 1830–2000* (Cambridge University Press, 2003)

A. Einstein, On the electrodynamics of moving bodies. Ann. Phys. **17**(10), 891–921 (1905)

D.H. Frisch, J.H. Smith, Measurement of the relativistic time dilation using muons. Am. J. Phys. **31**(5), 342–355 (1963)

J.C. Hafele, R.E. Keating, Around-the-world atomic clocks: observed relativist time gains. Science **177**(166), 168–170 (1972)

B. Hoffman, Noon-midnight red shift. Phys. Rev. **121**, 337 (1961)

Y.J. Huang, M. Fujieda, H. Takiguchi, W.-H. Tseng, H.-W. Tsao, Stability improvement of an operational two-way satellite time and frequency transfer system. Metrologia **53**, 881–890 (2016)

IERS Conventions (2010) Gérard Petit and Brian Luzum (eds.) (IERS technical note; no. 36), International earth rotation and reference systems service central bureau bundesamt für

kartographie und Geodäsie Richard-Strauss-Allee 11 60598 Frankfurt am Main Germany, available at https://www.iers.org/IERS/EN/DataProducts/Conventions/conventions.html
G. Kaplan, High-precision algorithms for astrometry: a comparison of two approaches. Astron. J. **115**, 361–372 (1998)
A.R. Klotz, Am. J. Phys. **83**, 231 (2015)
S. Kopeikin, M. Efroimsky, G. Kaplan, *Relativistic Celestial Mechanics of the Solar System* (Wiley-VCH Press, 2011)
A.P. Lightman, D.L. Lee, Restricted proof that the weak equivalence principle implies the Einstein equivalence principle. Phys. Rev. D **8**, 364–376 (1973)
D. Matsakis, Investigating a null test of the Einstein equivalence principle with clocks at different solar gravitational potentials, in *Proceedings of International Frequency Control Symposium*, New Orleans, La, 2016a
D. Matsakis, *Characterizing the Diurnal Signature in Two Way Satellite Time Transfer (TWSTT) Data with a Kalman Filter* (ION-PTTI, Monterey, 2016b)
S. Piel, S. Crane, J.L. Hanssen, T.B. Swanson, C.E. Ekstrom, Test of local position invariance using continuously running atomic clocks. Phys. Rev. A **87** (2013)
M. Rasel, *Quantum Tests of the Einstein Equivalence Principle on Ground and in Space* (IFCS-EFTF, Denver, 2015)
B. Rossi, D.B. Hall, Variation of the rate of decay of mesotrons with momentum. Phys. Rev. **59**(3), 223–228 (1941)
C. Will, The confrontation between general relativity and experiment. Living Rev. Relativ. **17**, 4 (2014)

Chapter 14
Real-Time Realization of UTC at Observatoire de Paris

G.D. Rovera, S. Bize, B. Chupin, J. Guéna, Ph. Laurent, P. Rosenbusch, P. Uhrich, and M. Abgrall

Abstract UTC(OP), the French national realization of the international Coordinated Universal Time, was redesigned and rebuilt. The first step was the implementation in October 2012 of a new algorithm based on an H-maser and on atomic fountain data. Thanks to the new implementation, the stability of UTC (OP) was dramatically improved and UTC(OP) competes with the best time scales available today. Then the hardware generating and distributing the UTC (OP) physical signals was replaced. Part of the new hardware is composed of commercial devices, but the key elements were specifically developed. One of them is a special switch that allows the UTC(OP) signals be derived from one of two time scales, based on two different H-masers, which are generated simultaneously. This insures the continuity of the UTC(OP) signal even when a change of the reference H-maser is required. With the new hardware implementation, UTC (OP) is made available through three coherent signals: 100 MHz, 10 MHz, and 1 PPS. For more than 3 years, UTC(OP) remained well below 10 ns close to UTC, with a difference even less than 5 ns if we except a short period around MJD 56650.

Keywords Coordinated Universal Time • Time scale • Hydrogen maser • Atomic fountain

Time and Time Scale

The word *time* is used to indicate two different concepts:

- The time interval, that is, the time elapsed between two events
- The proper time, that is, the position of a given event along a time line

G.D. Rovera (✉) • S. Bize • B. Chupin • J. Guéna • P. Laurent
P. Rosenbusch • P. Uhrich • M. Abgrall
LNE-SYRTE, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ, Paris 06, Paris, France
e-mail: daniele.rovera@obspm.fr

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_14

In both cases to exchange information, we must agree on the choice of a reference standard.

The time interval is one of the base units of the SI, defined as:

> The second is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium 133 atom.

The coordinate system almost universally accepted for event dating is UTC, at least for events occurring during our lifetime. Unfortunately there is no direct access to UTC, which is a paper time scale computed in different time. Therefore to date an event, it is necessary to measure the time difference between the event and a time marker of a physically available time scale. In everyday life the reading of a wristwatch is accurate enough to date events or to be in time to catch a train. The internal time scale of this clock is usually kept at a few seconds from UTC.

When the precise dating of an event becomes crucial, it is necessary to use a time scale tightly related to UTC. For this purpose in many countries, the National Metrology Institute (NMI) generates a time scale, available in real time, that is, the best realization of UTC for that country. This time scale denoted as UTC(k), where k represents the name of the laboratory, is usually the basis of the legal time for that country and serves as reference for the time scales at user level. Every month, the Bureau International des Poids et Mesures (BIPM) computes UTC taking into account hundreds of clocks all around the world and publishes the Circular T that, between other useful information, reports the time difference between UTC and the UTC(k) for the last completed month.

UTC(OP) is the time scale realized at Observatoire de Paris by the LNE-SYRTE that is the Designated Institute (DI) representing the French NMI LNE. Observatoire de Paris has a long history in time measurement and time scales. In 1676 Ole Roemer comparing the local time scale, based on Earth rotation, and a time scale based on Jupiter satellites, first demonstrated the finite speed of light.

In common meaning a UTC(k) is the physical realization of an *independent* prediction of UTC. The word *independent* is often forgotten and a number of UTC (k) is realized by using as reference some other UTC(k) or GPS Time with a short time constant phase lock. For many years UTC(OP) has been realized by a commercial Cs beam steered by a micro-phase stepper adjusted manually by an operator. Since October 2012 UTC(OP) is a true *independent* time scale based on H-maser automatically steered on purpose by a frequency offset generator driven by atomic fountain frequency data.

A detailed description of the UTC(OP) realization was recently published and is available in open access (Rovera et al. 2016); therefore, only recent results will be summarized here.

UTC(OP) Performances

The ultimate goal of UTC(OP) is the real-time realization of UTC with the best approximation. The indicator that better illustrates the obtained result is in this case simply the time difference between UTC and UTC(OP). This difference, since the

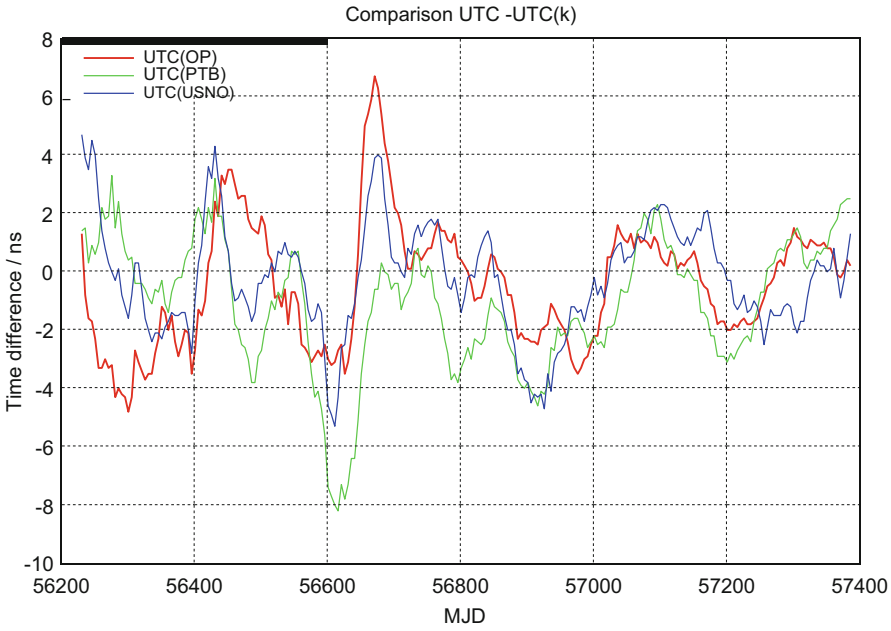


Fig. 14.1 Time differences between UTC and three fountain-based UTC(k) since the implementation of the new algorithm for UTC(OP)

implementation of our new algorithm, is reported in Fig. 14.1 together with the time difference between UTC and two other UTC(k) also using atomic fountains (Bauch et al. 2012; Peil et al. 2014). Since more than 3 years, UTC(OP) remains well below 10 ns close to UTC. The deviation is even less than 5 ns with exception of a short period around MJD 56600.

Conclusion

All the main hardware elements were progressively replaced with new ones, some of them designed on purpose. The transition phase is now achieved and UTC(OP) is currently the unique signal distributed to all time transfer equipment and to time users. All relevant intermediate signals have been characterized, as well as the UTC (OP) signals, and distributed at 100 MHz, 10 MHz, and 1 PPS. The UTC(OP) short-term noise is an exact replica of the H-maser noise for periods shorter than a few days. But for longer periods, it is much better than an H-maser.

UTC(OP) is now one of the best real-time realizations of UTC, thanks to the almost continuous operation of the LNE-SYRTE atomic fountains (Guéna et al. 2012), remaining well below 10 ns close to UTC. Since the beginning of 2015 UTC (OP), UTC is less than 3 ns. These performances reach the uncertainty of the

operational time transfer techniques that are currently the limiting factors in improving time scales.

Acknowledgments We gratefully acknowledge the support from the Service d'Electronique headed by Michel Lours and the technical staff of Reference Nationales de Temps. We acknowledge the useful discussion with Marco Siccardi and the realization of the frequency offset generators and of the switch by his company.

References

A. Bauch, S. Weyers, D. Piester, E. Staliuniene, W. Yang, Generation of UTC(PTB) as a fountain-clock based time scale. Metrologia **49**(3), 180 (2012)

J. Guéna, M. Abgrall, D. Rovera, P. Laurent, B. Chupin, M. Lours, G. Santarelli, P. Rosen-busch, M. Tobar, R. Li, K. Gibble, A. Clairon, S. Bize, Progress in atomic fountains at LNE-SYRTE. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **59**, 391–409 (2012)

S. Peil, J.L. Hanssen, T.B. Swanson, J. Taylor, C.R. Ekstrom, Evaluation of long term performance of continuously running atomic fountains. Metrologia **51**(3), 263 (2014)

G.D. Rovera, S. Bize, B. Chupin, J. Guéna, P. Laurent, P. Rosenbusch, P. Uhrich, M. Abgrall, UTC (OP) based on LNE-SYRTE atomic fountain primary frequency standards. Metrologia **53**(3), S81 (2016)

Chapter 15
Time in Television Systems

Donald Craig

Abstract Clocks are fundamental to the transmission of television—the when of the picture top and left hand side is as important for display as the image content, and maintaining audio lip sync has always been a battle. In overlapping periods since the modern television era began in 1941, there have been a variety of different clock synchronization and time labeling schemes. Real or imagined compatibility constraints during architectural transitions, as well as confusion of time the label versus time the physical property, have led to the evolution of remarkably byzantine clocks, with repeating decimal fractional frame rates, discontinuous timelines, and complex counting rhythms.

Beginning with monochrome transmissions, through the analog NTSC color system and into the digital ATSC system of today, this presentation explores the origins of the television system's clocks and views from the safety of hindsight unintended consequences of some innocent design optimizations. Clock system technical constraints included the collapse of logically independent layers for needed efficiency, synchronization of electromechanical as well as purely electronic systems, and transmission paths ranging from short-range coaxial cable to long-range variable length microwave RF. As a commercial and quasi-governmental enterprise, television also has had to accommodate the eccentricities of civil timekeeping, offering content to a schedule synchronized with wall clock time across multiple time zones.

Does it all need to be this complicated? A transition to television distribution over general purpose computer networks, coupled with the vastly improved performance ratio of low-cost modern digital circuitry, should eliminate the technical requirement for specialized clocks. The adoption of a new IEEE/SMPTE timing reference based on an epoch and an atomic time scale will provide an elegant mechanism for general purpose support of higher picture frame rates and audio sample rates. It remains for humans to abandon the old clocks and finally formalize the relationship with wall clock time in a way that doesn't unduly complicate system engineering.

D. Craig (✉)
Life Fellow, Society of Motion Picture and Television Engineers, Lexington, Massachusetts, USA
e-mail: dmc@arboretumstudios.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_15

Keywords Television • NTSC • ATSC • MPEG-2 Transport Stream • Frame rate • Synchronization • Audio/video synchronization • Lipsync • Timeline • Animation • Phenakistoscope • Remote electric clock • Raster • Interlace • Videotape • Timecode • Time as a label • Time as a physical property • Jam sync • Frame synchronizer • AES • SMPTE • IEEE • IEEE-1588 • SMPTE ST-12M • SMPTE ST-299 • SMPTE ST-2059 • AES-11 • Appointment TV • VOD • Joseph Plateau • Alexander Bain • Edwin Armstrong • David Sarnoff • William Paley

Chapter 16
From Computer Time to Legal Civil Time: IANA tz, IETF tzdist, etc.

Steve Allen

Abstract The internal time scale in computing and telecommunication systems is often based on UTC, but users of systems expect to see the local civil or legal time in their jurisdiction. Currently there are about 40–80 different time zones in active use. Historically there have been around 200 sets of rules needed to compute the differences between UT and local time. These rule sets have been painstakingly maintained by international agencies and corporations, and they are typically distributed to machines via updates to operating systems. Unfortunately, jurisdictions have changed their rules with little or no notice, and some machines do not get operating system updates in a timely fashion or ever. These lacunae are motivations for ongoing efforts among international standard organizations and corporations.

Keywords tzdist • Timezones • Operating systems • International Atomic Time (TAI) • Coordinated Universal Time (UTC) • Civil time • Legal time • Application programming interfaces (APIs)

The internal time scale in computing and telecommunication systems is often based on UTC, but users of systems expect to see the local civil or legal time in their jurisdiction. Currently there are about 40–80 different time zones in active use. Historically there have been around 200 sets of rules needed to compute the differences between UT and local time. These rule sets have been painstakingly maintained by international agencies and corporations, and they are typically distributed to machines via updates to operating systems. Unfortunately, jurisdictions have changed their rules with little or no notice, and some machines do not get operating system updates in a timely fashion or ever. These lacunae are motivations for ongoing efforts among international standard organizations and corporations.

S. Allen (✉)
University of California Observatories/Lick Observatory, ISB 260, 1156 High Street, Santa Cruz, CA 95064, USA
e-mail: sla@ucolick.org

E.F. Arias et al. (eds.), *The Science of Time 2016,* Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_16

Chapter 17
The UT1 and UTC Time Services Provided by the National Institute of Standards and Technology

Judah Levine

Abstract I will describe the network-based time services provided by the Time and Frequency Division of NIST. The network service is realized with approximately 20 servers located at many locations in the USA. The servers receive approximately 340,000 requests per second for time information in a number of different formats. The Network Time Protocol (NTP) receives approximately 98% of these requests. In addition to the servers that transmit UTC time, we also operate a single server that transmits UT1 time in NTP format. The UT1 server implements the offset between UT1 and UTC based on the data in the International Earth Rotation and Reference Systems Service (IERS) Bulletin A. The accuracy of the extrapolation is better than 4 ms; the accuracy of the time received by a user will depend on the stability of the network connection and is generally better than 8 ms. I will also describe the plan to improve the accuracy of the time service. This upgrade should be completed in the next few months.

Keywords Leap Second • NTP • TA(NIST) • UTC • UTC(NIST) • UT1

Introduction

I will describe the clock ensemble that is used to generate the NIST time scales UTC(NIST) and TA(NIST). The atomic time scale, TA(NIST), was synchronized to TAI in 1978 when it was first implemented as a real-time time scale. It has been free-running since that time. The UTC(NIST) time scale is realized as an offset from TA(NIST). It is steered toward UTC as computed by the International Bureau of Weights and Measures (BIPM). The steering is accomplished by automated frequency adjustments to the output phase stepper as I will describe below in more detail. The clocks are not steered.

J. Levine (✉)
Time and Frequency Division, NIST, Boulder, CO 80305, USA
e-mail: Judah.Levine@nist.gov

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_17

The digital time services provided by NIST respond to requests in several different formats that are received over dial-up telephone lines and the public Internet. The servers currently receive approximately 340,000 requests per second or about 3×10^{10} requests per day. About 98% of these requests are in the Network Time Protocol (NTP) format (Mills 2011; Mills 1991), and the remainder are in several other older formats, which are supported for legacy applications. The number of requests continues to increase at a rate of about 4% per month, and most of this growth is in NTP-format requests.

The NIST Time Scales

The NIST time scale (Levine 2012) uses an ensemble of commercial high-performance cesium standards and hydrogen masers. The input to the ensemble calculation is an array of time-difference measurements between one of the clocks in the ensemble, which is designated as the reference clock for the measurement process, and all of the other clocks. The time differences are measured every 12 min, and the time scale is calculated as soon as each measurement cycle is completed. The reference clock for the measurement process is chosen for its reliability and stability and has no special role in the ensemble calculation. The clock selected to be the reference clock also has no special role in generating the UTC(NIST) output signal.

The time-difference measurements are realized with a dual-mixer system (Stein et al. 1982). The output from each clock at 5 MHz is mixed with a frequency of 5 MHz–10 Hz, which is synthesized from the reference clock of the measurement process. The time difference of each clock is measured as the phase differences between the 10 Hz beat frequency generated by the mixing process and the 10 Hz frequency generated by the reference clock channel. The down-conversion of the 5 MHz input signals to 10 Hz increases the resolution of the measurement process by a factor of 5×10^5, so that a measurement resolution of 10^{-7} of a cycle at 10 Hz translates into an effective resolution of less than 1 ps at the input frequency of 5 MHz. The stability and accuracy of the time-difference data is limited by the corresponding parameters of the front-end mixer and is typically a few ps. This performance is routinely verified by measuring the time difference between a single clock connected to two measurement channels. The differential nature of the time-difference measurements attenuates any common-mode variation in the characteristics of the measurement channels.

The atomic time scale TA(NIST) is computed as the weighted average of these time differences. The weight assigned to each clock is derived from its stability, which is calculated as the RMS residuals of the time difference between the time of the clock after its deterministic parameters have been removed and the time of the ensemble averaged over the previous 30 days. This method overestimates the stability of a clock, since the clock is also a member of the ensemble that is used to evaluate its stability. This correlation is attenuated by decreasing the weight of a

high-weight clock (Tavella and Thomas 1991) and also by administratively limiting the weight of any clock to 30% of the ensemble. (Most time scale algorithms have similar limits.) Since the weights of all of the clocks must sum to 100%, limiting the weight of a good clock implicitly increases the weights of clocks that have poorer stability. Therefore, the administrative limitation on the maximum weight of any clock implicitly degrades the stability of the ensemble.

The frequency of TA(NIST) was set equal to the frequency of the International Atomic Time (TAI) frequency when the dual-mixer system was inaugurated, in 1978, and it has slowly drifted upward in frequency since that time. The current frequency of TA(NIST) is greater than the frequency of TAI by approximately 37 ns/day (a fractional frequency of approximately 4.5×10^{-13}).

The UTC(NIST) time scale is derived from TA(NIST) by an offset in time and frequency that is calculated administratively. The offset is a piecewise linear function that has no time steps. That is, changes to the offset equation are realized as changes in frequency only. The origin time of each of the linear steering equations is the ending time of the previous function. In other words, the origin time of each steering equation is the integral of all of the previous frequency offsets from 1978 when the system was initialized to the current time. The steering frequency has been negative for many years to remove the positive frequency offset of TA(NIST) that was described in the previous paragraph. The current and proposed future steering equations are published in the periodic Bulletin of the NIST Time and Frequency Division (NIST 2017a).

The steering adjustments are computed from the data in the BIPM Circular T (BIPM 2017a), which gives the difference between the UTC time scale computed by the BIPM and the time scales of all of the contributing laboratories (BIPM 2017b). The Circular T for any month is typically published about the 10th day of the following month. The BIPM also computes a more rapid version of UTC. The rapid version, UTCr (BIPM 2017c), provides daily estimates of the same data as in Circular T with a delay of 1 week. The results of the UTCr calculation are published on Wednesday afternoon (about 1700 UTC) for the period ending on the previous Monday. Starting in January, 2016, the data from UTCr are also used to compute a steering adjustment for UTC(NIST). This is done to improve the short-term accuracy of the UTC(NIST) time scale. The BIPM data for UTC-UTC(NIST) and UTCr-UTC(NIST) are shown in Fig. 17.1.

After each 12-min time scale computation, the software computes the offset in time and in frequency of UTC(NIST) with respect to TA(NIST) based on the appropriate steering equation and applies these parameters to a hardware phase stepper that produces the requested output signals at 5 MHz and 1 Hz. The reference signal for the phase stepper is derived from one of the clocks in the ensemble. The signal from the reference clock itself is not steered. The reference clock for the phase stepper has no special significance in the time scale algorithm, and is chosen mostly for its reliability and longevity. In general, the reference signal for the phase stepper is derived from one of the hydrogen masers in the ensemble, because these signals have the best short term frequency stability. The 5 MHz output signal from the phase stepper is fed back into the time scale measurement system, where it is

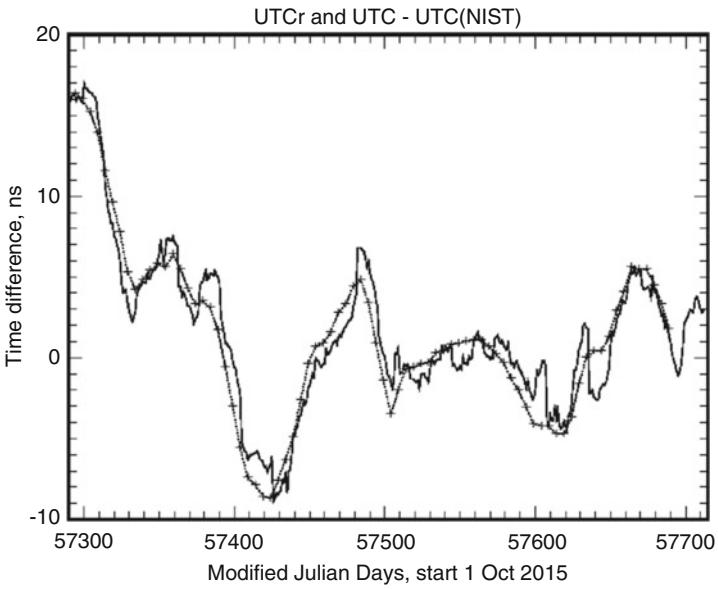


Fig. 17.1 The time differences in ns between the UTC and rapid UTCr time scales computed by the BIPM and UTC(NIST). The *dotted line* with the symbols shows the UTC time difference, and the *solid line* is the UTCr time difference. The *X* axis is in Modified Julian Days, and the plot shows the time differences from October 2015 through November 2016

measured along with all of the other clocks in the ensemble. However, the signal from the phase stepper is not included in the ensemble calculation, and its time difference with respect to the reference clock of the measurement system is used to verify that the requested steering offsets have been correctly applied. In normal operation, the difference between the physical output of the phase stepper and the equation that defines UTC(NIST) for that epoch is a few ps and is well characterized as white phase noise.

The output signals from the phase stepper are used as the reference signal for all of the NIST time services. They are also used to control the link between NIST and the BIPM through PTB (the Physikalisch-Technische Bundesanstalt, which is the National Metrology Institute of Germany). This link is used by the BIPM to compute the values for UTC-UTC(NIST) that are reported in the monthly Circular T and in the more rapid UTCr data. The link between NIST and PTB is implemented using the two-way message exchange protocol, and this method is also used to support the NIST digital time services. I will describe this method in the next section.

The Two-Way Method

Estimating the delay in transmitting messages from the source to the user is usually the most important and most difficult aspect of any time distribution system. The two-way method, in which the one-way delay is estimated as one-half of a measurement of the round-trip value, is widely used for this purpose.

There are a number of different realizations of the two-way method. In the version that is generally used to measure the delay in packet-switched networks, the local station transmits a time query to the remote system, and the remote system replies. The systems at both ends transmit continuously in the satellite implementation of the protocol. In both cases, the two transmission time stamps and the two received time stamps are combined to estimate the transmission delay and the time difference between the clocks at the two stations. The details of the calculation are described in the following paragraph.

Let T_{1s} be the time when system 1 transmits a message to system 2 as measured by the clock in system 1. System 2 receives the message at time T_{2r} and replies at time T_{2s}, where both of these times are measured by the clock in system 2. This second message is received back at system 1 at time T_{1r}, again measured by the clock in system 1. The round-trip travel time, as measured by system 1, is given by

$$\Delta = (T_{1r} - T_{1s}) - (T_{2s} - T_{2r}) \tag{17.1}$$

The first term is the total elapsed time for the message exchange, and the second term is the delay between when the second system received the message and when it replied. For systems that transmit continuously, the messages in both directions are combined so that the second term is positive. If the path delay in one direction is d, then the message transmitted at time T_{1s} arrives at the second system at time $T_{1s} + d$ as measured by the clock in system 1. The time of system 2 at that instant is T_{2r}, so that the time difference between systems 1 and 2 is estimated by

$$\delta t_{12} = (T_{1s} + d) - T_{2r} \tag{17.2}$$

If the path delay is symmetrical, then $d = \Delta/2$, and the time difference between systems 1 and 2 is given by

$$\delta t_{12} = \frac{T_{1s} + T_{1r}}{2} - \frac{T_{2s} + T_{2r}}{2} \tag{17.3}$$

The accuracy of the two-way method depends on the assumption that the delay is symmetric—that the inbound and outbound delays are equal. The inbound and outbound delays are comprised of two contributions: the delays through the station hardware at each end point and the delay through the channel itself. The delays through the station hardware can be measured in principle, and any asymmetries can be calibrated and removed administratively. The uncorrelated variations in the inbound and outbound delays in the station hardware must be minimized as much as

possible. A residual admittance to fluctuations in the ambient temperature is a common problem, and variations in the ambient temperature must be minimized for that reason.

The symmetry of the channel delay is a more difficult problem, because it is usually not under the control of the operators at either station. The symmetry requirement is not too difficult to realize with simple physical circuits, but it is much more difficult to do so when the circuit is realized as a packet-based communications channel with intervening routers, switches, and gateways. The queueing delays in these elements often depend on the traffic load, especially when that load is asymmetric, as is often the case with systems that provide web-based services or streaming audio or video information. Even under the best of circumstances, an asymmetry of a few percent of the total round-trip delay is quite common in packet-switched networks, and this asymmetry often sets the limit on the accuracy of any distribution method that uses packet-switched networks such as the public Internet. Since the round-trip delay is often on the order of 100 ms in a packet-switched network, these networks typically provide accuracy that is often not much better than a few milliseconds. This limit has nothing to do with the format of the messages or the protocol that is used to transfer the information, and there is no advantage of one protocol over another to the extent that the accuracy is limited by this consideration.

It is possible to mitigate the impact of the asymmetric delay contributions of routers, switches, and gateways in the communications channel by special purpose hardware at each one of these network elements (IEEE 2008). The special-purpose hardware bridges the network element and measures the latency delay of a message passing through the device. The details of how the latency is measured vary somewhat from one implementation to another, but the important point is that this technique removes one of the primary limitations to the accuracy of packet-based time distribution. Unfortunately, these extra devices are not present in the general public Internet, so that the improvement in accuracy is not usually possible when the public Internet is used for time distribution.

Although the one-way delay in a link between two ground stations that uses a synchronous communications satellite link is quite large, the asymmetry is very small, so that time comparisons that use this method can realize sub-microsecond accuracy. The stability and reciprocity of the delay through the satellite transponder and the ground station equipment often limit the accuracy of the comparisons.

The NIST Time Servers

The Time and Frequency Division of NIST operates a number of network-based time servers that are synchronized to the NIST clock ensemble. The servers that are located at the NIST facilities in Colorado are synchronized by a direct hard-wired connection between the 1 pulse-per-second signal that realizes UTC(NIST) and the server hardware. The accuracy of the time at the server is a few microseconds. The

servers that are located at other locations (see the Internet Time Service list at http://tf.nist.gov for a list of the locations) are synchronized by a telephone connection between the server and the NIST clock ensemble in Boulder. The connection uses the ACTS protocol (Levine et al. 1989), which transmits time over dial-up telephone lines. The delay in the telephone connection is measured by using the two-way protocol described above, and the accuracy of the time at the server is typically a few milliseconds.

The servers respond to time requests received over the public Internet in a number of different formats. The most common request is in the Network Time Protocol (NTP) format. This format implements the two-way method for estimating the transmission delay, and the accuracy of the messages is limited by the considerations that I discussed in the previous section.

There are three types of NTP servers. Most of the NIST servers are synchronized to UTC(NIST) and respond to requests by using the standard NTP message format. The second type of server also transmits time data based on UTC(NIST) but includes support for message authentication based on the MD5 (Rivest 2011) or SHA1 (National Institute of Standards and Technology 2015) hash codes. There are currently approximately 400 users of this service, and each one has a unique key. The key is combined with the standard NTP request and reply packets, and the hash function of the combination is computed by the sender and appended to the message. The receiver computes the hash function of the message and compares the computed result with the transmitted version. The message is accepted only if the computed hash function agrees with the one in the message. Since the key parameter is a secret known only to NIST and the single registered user, the authentication process ensures that the inbound and outbound messages were not sent by a rogue third party and were not modified in transmit. The authentication process does not affect the accuracy of the time transmitted in this way, since the accuracy is limited by symmetry of the path and is not affected by the contents of the message.

The third type of NTP server transmits UT1 time in the standard NTP format. The UT1 time is computed by adding the difference between UT1 and UTC published by the International Earth Rotation and Reference Service (IERS) in Bulletin A (IERS 2017). The bulletin is published every week and contains estimates of the UT1-UTC time difference for about a year in the future. The accuracy of the prediction for 30 days in the future is approximately 4 ms, and this is comparable to the accuracy of the NTP time transfer process as I have described above. Therefore, the values used by the time server are updated every few weeks, and the longer-term predictions in each bulletin are never used.

There are currently 24 registered users of this service. The service was originally limited to registered users so that the general user, who may not understand the difference between UTC and UT1, will not connect to the server by mistake and receive a time message that can be "wrong" by up to 0.9 s. The server was converted to open access in September 2016. It currently receives requests from about 900 unique network addresses.

In addition to responding to requests in the NTP format, the servers also respond to requests in the DAYTIME (TFC 1983) and TIME protocols (Postel 1983). These are simpler protocols that do not support the full two-way message exchange, and they are used when the client hardware is relatively simple and when the highest accuracy is not required. The NIST implementation of the DAYTIME protocol includes advance notice of the transitions to and from daylight saving time based on the US model and also provides advance notice of leap seconds. The daylight saving time information is particularly useful for older systems that may not have any other source for this information or that may have the older, incorrect US transition dates. The TIME protocol is the simplest of all of the message formats and is supported for legacy applications only. We do not recommend it for new applications because it has poor error-checking capabilities, supports no method for estimating the network delay, and provides none of the ancillary information of the DAYTIME format.

The NIST time servers receive approximately 340,000 requests per second for time in the various formats, and almost all of these requests (>98%) are in the NTP format. The number of requests has been increasing at a rate of approximately 4% per month, and this rate of increase shows no sign of moderating. The increase in the number of requests since 2004 is shown in Fig. 17.2. The figure shows that the number of requests in the DAYTIME and TIME formats has been decreasing since 2012, but there are still approximately 10^8 requests per day in these formats or about 1000 requests per second.

Traceability

The concept of traceability is important for all time services. A time measurement is traceable if there is an unbroken chain of measurements from the end user back to a national or international standard of time (Taylor and Kuyatt 1994). Each link in the measurement chain must be characterized by a delay and an uncertainty. In the USA, traceability can be satisfied by a chain that leads back to UTC(NIST) or UTC (USNO).

In addition to the technical traceability requirement of the previous paragraph, some users are required to have legal traceability. In addition to the technical requirements for traceability, legal traceability adds requirements that can confirm technical traceability in a legal adversary proceeding. The details of how legal traceability is satisfied depend on the end user but generally include properly maintained and certified log files and other documentations that can confirm the proper operation of the synchronization process. A log file that contains only error information may be inadequate, because it can be completely empty when the system is working properly, and it can be hard to distinguish this case from a system that is totally inoperative.

Although a time signal may be generated by a traceable source such as a GPS satellite or a NIST-operated time server, the traceability of these signals generally

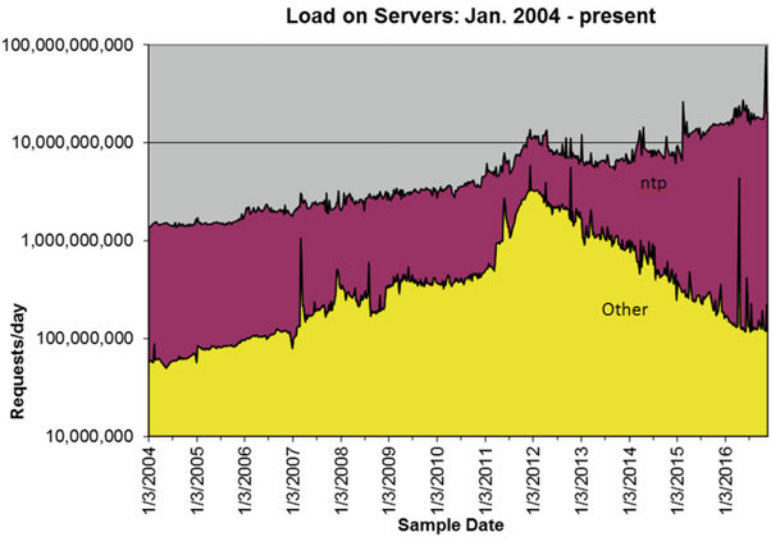


Fig. 17.2 The number of time requests received by all of the NIST time servers from January 2004 to November 2016. The logarithm of the number of requests as a function of time is displayed. The *lower curve*, labeled "other," is the number of requests in the TIME and DAYTIME formats, while the *upper curve* is the total number of requests, so that the difference is the number of requests in NTP format

does not extend to un-calibrated and unmonitored user equipment. That is, the traceability of a signal from a GPS satellite ends with the signal in space by default, and the default traceability of the time from a NIST time server ends at the server portal. The messages transmitted by third-party servers are generally not traceable.

The Leap Second Problem

The current definitions of the length of the second in both the UTC and TAI time scales are based on the same value derived from the frequency of a hyperfine transition in the ground state of cesium 133 (BIPM 2006). The length of the second that results from this definition is somewhat shorter than the length of the second of the UT1 time scale, which is derived from the rate of rotation of the Earth. The difference between the two time scales has a systematic contribution whose magnitude is currently somewhat less than 1 s per year. This systematic increase is caused by tidal friction and a number of other, smaller effects. (Nelson et al. 2001) There is also an additional contribution that is irregular and whose magnitude cannot be accurately characterized either deterministically or stochastically. The stochastic contribution to the length of the UT1 second could reverse the sign of the UTC-UT1 difference in principle, but this has not happened to date and is unlikely to happen in the future. The UTC time scale is used as a proxy for UT1 time in many

applications ranging from every day timekeeping to the timing of events in astronomy, and maintaining the equivalence between UT1 and UTC is an important consideration in timekeeping.

In order to prevent the UTC and UT1 time scales from diverging, an extra "leap" second is added to UTC whenever the magnitude of the difference between UTC and UT1 approaches 0.9 s (NIST 2017b). (As I mentioned in the previous paragraph, it could be necessary in principle to drop a second in UTC to preserve this equivalence, but this is unlikely in the foreseeable future.) This extra second is added after 23:59:59 UTC, usually on the last day of June or December. (The last days of other months are possible in principle but have never been used.) The name of this extra second is 23:59:60 UTC. The leap second is defined with respect to UTC, and it therefore occurs at different local times at different locations. For example, it occurs late in the afternoon in California and during the following morning in Asia and Australia. The second after the leap second is named 00:00:00 UTC of the next day.

Computer clocks and most digital systems keep time internally as the number of seconds that have elapsed since an origin time, which can differ from one system to another. The conversion between the count of the elapsed seconds and the time in the conventional year-month-day hour-minute-second representation is computed by the operating software whenever the time is requested. Additional offsets for the extra day associated with leap years, for conversion to the local time zone, and for daylight saving time are also added if necessary. The various offsets are either strict constants (as for the offset of the local time zone) or can be computed algorithmically (as for the offsets for a leap year or for the local version of daylight saving time). Both leap days and the daylight saving time transitions introduce discontinuities between a time interval computed from the count of the elapsed seconds maintained by the system internally and a simple computation of the same time interval using the hour-minute-second representation. This discontinuity can be particularly serious during the Fall transition away from daylight saving time, since the local time moves backward by 1 h during the transition and the same time stamps will be transmitted twice. There is often no indication whether the time reported by the system during this transition period is the old daylight time or the new standard time. The impact of these changes is mitigated by the fact that the change is made at 2 a.m. Sunday local time.

The method of timekeeping in a digital system is therefore divided into two cooperating processes: a low-level application that receives and counts the periodic signals generated by the hardware and a second higher-level application that "knows" how to convert this count to the conventional representation of hours-minutes-seconds. In general, the partition between the low-level and high-level processes is arbitrary because the application need not be concerned with the intermediate values transmitted between them. However, leap seconds are inserted on an irregular and unpredictable basis, and the times of both future and past leap seconds cannot be computed algorithmically. Therefore, there is no simple way of incorporating leap seconds into this deterministic pair of cooperating processes.

The solution used by the NIST time servers and by many other systems is to stop the hardware clock for 1 s at 23:59:59 UTC on the day of a leap second. The result is to use the time 23:59:59 UTC twice—once when that time occurs and a second time during the leap second. Neither the system applications nor the standard communications protocols have any provision for indicating that the second 23:59:59 time value is actually a leap second and should be assigned the time 23:59:60. This solution has a number of undesirable side effects. The fact that there are 2 s with the same name introduces confusion and can produce time stamps that appear to violate causality: 23:59:59.2 during the leap second actually occurred after 23:59:59.5 during the first second with this time stamp value, but the time stamps themselves indicate the opposite order. (The Fall transition from daylight saving time to standard time has the same problem with time stamps between 2 a.m. and 3 a.m.)

This method of repeating 23:59:59 is probably the closest approximation to the definition of the leap second on systems that have no method of indicating the leap second event to applications and inserting a special leap second indicator into communications protocols. The NIST DAYTIME protocol transmits time in the hour-minute-second format and transmits the leap second time correctly as 23:59:60.

Although the method I have just described has undesirable side effects, there are other implementations that are worse. For example, a number of network time-service appliances add the leap second to the first second of the new day. That is, the time stamps in the vicinity of a leap second are 23:59:58, 23:59:59, 00:00:00, and 00:00:00. The long-term behavior of this method is correct, but it adds the leap second to the wrong day and has a transient error of 1 s with respect to the definition. Other systems implement the leap second as a frequency adjustment over some time interval before (or after) the leap second (https://googleblog.blogspot.com/2011/09/time-technology-and-leaping-seconds.html). This method introduces a varying time error on the order of 0.5 s with respect to the definition and a frequency error that depends on the interval over which the leap second is inserted. Since the interval for amortizing the leap second is not defined in any standard, different implementations of this method may produce time stamps in the vicinity of a leap second that disagree both with each other and with respect to UTC.

Finally, it is clear that all of these solutions introduce steps in both time interval and frequency in the vicinity of a leap second and that a real-time physical process does not stop during the leap second. The time scales of navigation systems do not use leap seconds for this reason, although each navigation system time typically includes the number of leap seconds that have been inserted from the start of the leap second system in 1972 to the instant that the navigation system time scale was initialized. Each navigation system time will therefore have a different offset from UTC for this reason, and all of these offsets will change whenever a new leap second is inserted. This multiplicity of time offsets from UTC is undesirable because it can be a source of confusion when time stamps are derived from the transmissions of navigation satellites without a clear indication of which time scale was used.

A number of proposals have been advanced for addressing the leap second problem, but a detailed discussion of them is outside of the scope of this document, which is focused on the NIST time services.

Improvements to the Service

The Time and Frequency Division of NIST is addressing the continued increase in the number of requests for time services by upgrading the capacity of the servers and the network connections. We plan to have a number of very high capacity sites operating in the next few months. In addition, we are adding reference clocks to each of these sites to improve the accuracy of the time at the servers. We expect that the reference clocks at the remote sites will be within 10–20 ns of UTC(NIST), and we will consider offering services that can transmit the time to a user's system with an end-point accuracy of 100 ns or better.

Summary and Conclusions

The NIST Time and Frequency Division operates an ensemble of time servers that respond to requests in a number of different standard formats. The servers receive approximately 340,000 requests per second, most of which are in the NTP format. The time of the servers is directly traceable to the UTC(NIST) time scale, which is maintained by an ensemble of atomic clocks and hydrogen masers located at the NIST laboratory in Boulder, Colorado.

NIST operates a single server that transmits UT1 time in the standard NTP format. The internal time of the server is synchronized to UTC(NIST), and the difference between UT1 and UTC is added to the NTP response just before it is transmitted to the system that requested it. The UT1-UTC time difference is derived from the International Earth Rotation and Reference System Service (IERS) Bulletin A, and the value of the UT1-UTC difference added to each message is calculated by a simple linear interpolation of the previous and following values in Bulletin A. The server is open access and currently receives requests from about 900 unique network addresses.

NIST also operates three time servers that authenticate the time stamps by computing the MD5 or SHA1 hash of the combination of a secret key and the standard time message. Each registered user of the authenticated service is given a unique key. There are currently approximately 400 users of this service. The authentication process ensures that the message originated from NIST and was not modified in transit. It does not affect the accuracy of the time service, which is limited by the symmetry of the network connection between the client and the server.

The time servers implement a leap second by transmitting the time corresponding to 23:59:59 a second time during the leap second event. This method is used because neither the system nor the NTP message format can represent the correct name of the leap second, which is 23:59:60. The DAYTIME format uses the hour-minute-second representation and transmits the correct representation of the leap second. There is no discontinuity in the time transmitted by the UT1 time server across the leap second.

References

BIPM, *Le Systeme International d'Unites*, 8th edn. (BIPM, Sevres, 2006)

BIPM, Published monthly by the International Bureau of weights and measures and available from the BIPM web site at ftp://ftp2.bipm.org/pub/tai/publication/cirt/ (2017a)

BIPM, Details and data for all contributing laboratories are available on the BIPM web page at www.bipm.org (2017b)

BIPM, The weekly UTCr data for all of the contributing laboratories are published in tables on the BIPM web page at: ftp://ftp2.bipm.org/pub/tai/publication/utcr. All of the data for each contributing laboratory is at: ftp://ftp2.bipm.org/pub/tai/publication/utclab (2017c)

IEEE, IEEE Standard for a Precision clock synchronization protocol for networked measurement and control systems, IEEE Standard 1588, available from: webstore.ansi.org. 2008

IERS, International Earth Rotation and Reference Service, http://datacenter.iers.org/eop/-/somos/5Rgv/latest/6 (2017)

J. Levine, The statistical modeling of atomic clocks and the design of time scales. Rev. Sci. Instrum. **83**, 012201–012228 (2012)

J. Levine, M. Weiss, D.D. Davis, D.W. Allan, D.B. Sullivan, The NIST automated computer time service. J. Res. NIST **94**, 311–322 (1989)

D.L. Mills, RFC 1305, available from: www.ietf.org/rfc/rfc1305.txt. 1991

D.L. Mills, *Computer Network Time Synchronization*, 2nd edn. (CRC Press, Boca Raton, 2011), p. 64

National Institute of Standards and Technology, Federal Information Processing Standard FIPS Pub 180-4, 2015

R.A. Nelson, D.D. McCarthy, S. Malys, J. Levine, B. Guinot, H.F. Fliegel, R.L. Beard, T.R. Bartholomew, The leap second: Its history and possible future. Metrologia **38**, 509–529 (2001)

NIST, The Steering data is available online at http://www.nist.gov/pml/div688/grp50/nistat1.cfm (2017a)

NIST, Leap Seconds are described at www.nist.gov/pml/div688/leapseconds.cfm. See also: www.itu.int/en/ITU-R/Pages/default.aspx (2017b)

J. Postel, The Time Protocol, RFC 866, www.rfc-editor.org/info/rfc866, 1983

R. Rivest, The MD5 Message-Digest Algorithm, Request for comments RFC-1321, 1992. Updated in request for comments RFC-6151, 2011. The documents are available from www.ietf.org/rfc, 2011

S. Stein, D. Glaze, J. Levine, J. Gray, D. Hilliard, D. Howe, L. Erb, Performance of an automated high accuracy phase measurement system, in *Proceedings of the 36th Annual Symposium on Frequency Control*, Fort Monmouth, New Jersey, 1982. Reprinted in NIST Technical Note 1337, available from the publications list at tf.nist.gov

P. Tavella, C. Thomas, Comparative study of time scale algorithms. Metrologia **28**, 57–63 (1991)

B.N. Taylor, C.E. Kuyatt, *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results, NIST Technical Note 1297* (U. S. Government Printing Office, Washington, D.C., 1994)

TFC, The NIST version of the DAYTIME protocol is described on internet time service page at: tf.nist.gov. See also, J. Postel, www.tfc-editor.org/info/rfc867, 1983

Chapter 18
On a Redefinition of the SI Second

Fritz Riehle

Abstract Optical atomic clocks like optical lattice clocks or single-ion clocks have outperformed the best cesium atomic clocks that realize the base unit of time in the International System of Units (SI) with respect to accuracy and stability. With fractional uncertainties in the 10^{-18} regime and fractional instabilities of a few 10^{-16} in 1 s, optical atomic clocks are presently the most accurate measuring devices at all.

The status of optical clocks and the different means to compare remote clocks will be reviewed. Recent experiments to address the question of the constancy of fundamental constants or their application in the novel field of relativistic geodesy will be discussed. All these developments will eventually ask for a new definition of the unit of time based on optical frequency standards, thereby replacing the cesium standard in the SI. In the second part of this talk, the question what will be the necessary prerequisites and challenges for such a redefinition will be addressed. The current status of the discussion will be reported and an attempt for a tentative roadmap toward such a redefinition of the second in the SI will be given.

Keywords Optical atomic clocks • Redefinition of the second • Relativistic geodesy • Constancy of fundamental constants

F. Riehle (✉)
Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
e-mail: fritz.riehle@ptb.de

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_18

Chapter 19
Time Scales Steered by Optical Clocks

T. Ido, H. Hachisu, F. Nakagawa, and Y. Hanado

Abstract Although optical frequency standards made rapid progress these days, microwave standards are still employed as source oscillators of time scales because an oscillator free from phase jumps is a prerequisite. Currently, for high-end users such as national metrology laboratories, hydrogen masers (H-masers) have adequate balance of the stability and reliability. Thus, optical clocks may play the role of the standards to which the time scales refer to in order to adjust their scale intervals. The benefit of using optical frequency standards, in this case, would be the capability to evaluate the scale interval of the H-maser more quickly. To investigate such possibilities of "optical" steering, we evaluated the behavior of an H-maser over a few months with reference to a ^{87}Sr lattice clock. The evaluations clearly demonstrated a stable linear drift of the H-maser frequency, indicating the capability of compensating the drift for a stable time scale. This prospect was also supported by a numerical simulation based on the record of H-maser-UTC (NICT)-UTC link.

Keywords Time scale • Lattice clock • TAI • UTC • Steering

Introduction

The recent progress of optical clocks is so rapid that some atomic clocks based on optical transitions have clearly surpassed the microwave clocks in terms of accuracy and stability. Hence, optical clocks may be utilized for maintaining a time scale. Time scales have so far relied on cesium fountain frequency standards for the accuracy of the scale interval. Particularly, some laboratories operate fountains almost continuously and adjust their time scales in real time according to the error that fountains detect (Bauch et al. 2012; Abgrall et al. 2015; Peil et al. 2016). For laboratories without Cs fountains, on the other hand, it is not easy to begin the

T. Ido (✉) • H. Hachisu • F. Nakagawa • Y. Hanado
Space-Time Standards Laboratory, National Institute of Information and Communications Technology (NICT), Koganei, Tokyo 184-8795, Japan
e-mail: ido@nict.go.jp

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_19

development from scratch only for the purpose of time scale maintenance. Thus, in the near future, optical clocks will be required to play a role in maintaining a time scale.

UTC: Reference Time Scale to Which Local Time Scales Synchronize

Coordinated Universal Time (UTC) is one of the most stable time scales and is referenced by almost all national standard times. The scale interval of UTC is same as that of International Atomic Time (TAI), which is based on the average of more than 500 commercial atomic clocks operated in various metrological or astronomical institutes worldwide. The interval obtained by the mean is adjusted to the SI second by referring to the calibrations provided mainly by cesium fountain standards operated in national metrological laboratories. The averaging and adjustment are virtually performed by a numerical process. Thus, there is no real signal of UTC. On the other hand, participating institutes normally maintain their time scales, generating real-time clock signal, and are adjusted as close as possible to UTC. These local realizations of UTC are designated by UTC(k), where k denotes the name of the institute. The Bureau International des Poids et Mesures (BIPM) is in charge of the numerical process of calculating TAI (http://www.bipm.org/en/bipm/tai/tai.html). The "tick" of TAI is once in 5 days since this numerical process is performed for the dates at 0:00 UTC on every 5 days. The BIPM notifies the participating institutes of the time difference between UTC and UTC(k). The notification is performed in a monthly report, the so-called Circular T. The institutes steer the scale interval according to this notification. Note that the BIPM also publishes a rapid UTC (UTCr), where the time differences are reported daily. However, the shorter average of a single day suffers from rather large link uncertainty (http://www.bipm.org/en/bipm-services/timescales/rapid-utc.html).

The difficulties of this steering are partly due to the latency of the Circular T. The time-keeping institutes are unable to obtain the time difference UTC − UTC(k) in real time. In other words, for the duration of one month, the scale interval of UTC (k) is basically determined by their local frequency standards. Hence, locally available atomic frequency standards are important for the time scale. Cesium fountain frequency standards were first developed in the 1990s, and now the operation of some standards in advanced laboratories is so stable that an up-time ratio of more than 90% is maintained (Bauch et al. 2012). Thus, some institutes can maintain accurate scale interval of the time scale in real time (Bauch et al. 2012; Abgrall et al. 2015). This improvement is clearly witnessed in the record of UTC − UTC(k) in the past ten years, which is shown as Fig. 19.1. Reduction of the fluctuations in UTC − UTC(PTB) and UTC − UTC(OP) is found at around MJD 55500 and MJD 56200, respectively, where the two institutes initiated the steering of UTC(k) by referring to real-time evaluation using a reliably operated Cs fountain.

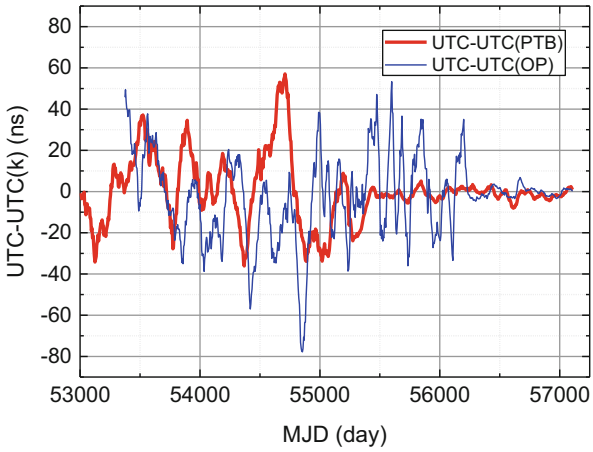


Fig. 19.1 Time difference between UTC and two local realizations UTC(k)

Absolute Frequency Measurement Based on TAI

For the steering of time scales using optical clocks, the absolute frequency of clock transitions needs to be rigorously evaluated to the limit of the SI second. It is straightforward to measure the optical frequency by referring to locally available cesium primary frequency standards. The evaluation of the frequency based on TAI, however, is an alternative for laboratories without highly accurate primary frequency standards.

Every month, the BIPM reports in Circular T the result of the calculation of UTC − UTC(k). Consequently, the clock frequency can be evaluated only for the frequency of a 5-day average. Thus, it is ideal to operate the clocks under test continuously for 5 days over one TAI grid. Instead of the continuous operation, however, we employed a local transfer oscillator (LTO) and measured the frequency with reference to the 5-day average of the LTO, which connects to UTC via UTC(k) in the 5-day-mean basis.

The absolute frequency of a ^{87}Sr optical lattice clock in NICT was first evaluated in 2012 based on TAI using a scheme as shown in Fig. 19.2a, with an uncertainty of 3.3×10^{-15} (Yamaguchi et al. 2012). UTC(NICT) adjusts the frequency of the source oscillator every 8 h by referring to the Cs ensemble time, which may cause sporadic frequency changes of 2×10^{-15} at maximum. Since no care was taken to ensure the homogeneous distribution of measurement time over the 5-day grid of TAI, a large dead time uncertainty of 2.6×10^{-15} was required. It is crucial to reduce the TAI link uncertainty in order to evaluate the absolute frequency more accurately. Thus, we performed absolute frequency measurements of the ^{87}Sr lattice clock extensively as shown in Fig. 19.2b (Hachisu and Ido 2015). In this case, the measurement is affected not by the fluctuation of the $y_{\mathrm{UTC(k)}}$ but is determined by the time difference LTO − UTC(k) and UTC(k) − UTC at day 0 and day 5, which can be obtained with reduced uncertainties.

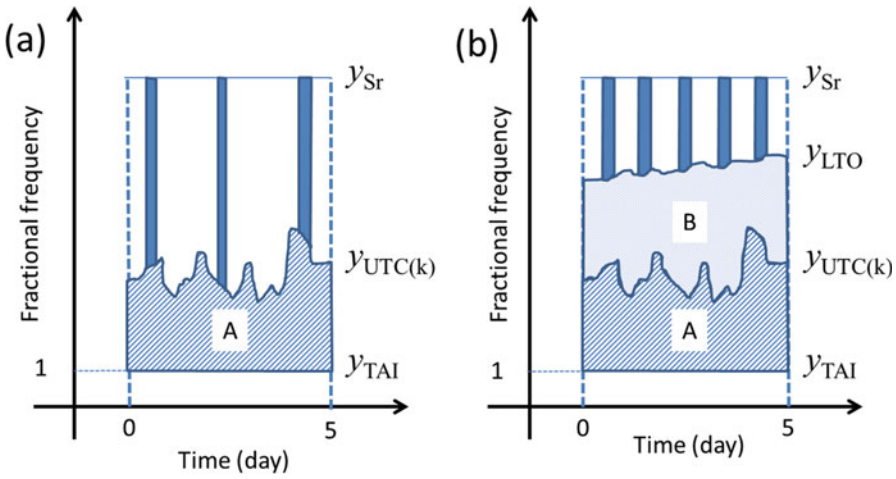


Fig. 19.2 Schematic diagram for an absolute frequency measurement using TAI. (**a**) The time differences between UTC(k) and UTC at 0:00 (UTC) of day 0 and day 5 are reported by Circular T issued by the BIPM, which corresponds to area A. (**b**) A highly stable local transfer oscillator (LTO) reduces the ambiguity of the LTO frequency. Area B is obtained by a time interval counter

Three sets of the 5-day frequency measurements were performed in Hachisu and Ido (2015). The weighted average of the three campaigns was calculated on the basis of the three means according to a statistical weighting. Combining the statistical uncertainty, the systematic uncertainties of the clock, link uncertainties, and the dead time uncertainties which are estimated by numerical analysis, the absolute frequency of the ^{87}Sr lattice clock transition was evaluated to be 429,228,004,229,872.85(47) Hz.

Capability of Keeping a Time Scale by the Complementary Contributions of an Optical Clock and a Highly Predictable Source Oscillator

In our recent campaign of absolute frequency measurements (Hachisu and Ido 2015), measurements based on an H-maser (HM4) were performed in four campaigns which are temporarily separated by 20–30 days. Conversely, this measurement allows us to obtain the scale interval of the HM4 based on the Sr lattice clock. A sample of the evaluation of HM4 frequency by the ^{87}Sr lattice clock is shown in Fig. 19.3. Four campaigns were performed on MJD 57060–57064, 57079–57084, 57108, and 57124–57129. The calibrations were expressed as four groups of green dots. All points are derived from the calibration of 10^3–10^4 s. The type-A uncertainty of each point is $<1 \times 10^{-15}$. Whole points may have an offset of $<5 \times 10^{-16}$ due to the error in assuming the ^{87}Sr clock frequency. Overall points indicate a

fairly stable linear frequency drift of HM4. This was also confirmed by the instability against UTC, which is shown as the inset of Fig. 19.3. Thus, we can estimate the frequency of HM4 on any instant by interpolating past data by a linear regression. The thick line indicates such estimation. Using the data taken in the last 45 days, the linear drift of the H-maser is estimated, and we interpolate it to estimate the HM4 frequency on that date. Since the data of the first 4 days (MJD 57060–57064) is insufficient to estimate the overall drift rate of the HM4, the red thick line is probably steeper than the real drift rate for the first 30 days. However, the data on MJD 57079, which is a point after 15 days of down time, allows the system to notice that the estimation of the drift rate was higher. Later than that, the estimation of the HM4 frequency becomes fairly good. The blue thin line is another estimation of the HM4 frequency provided by the JST (Japan Standard Time) system. JST is a product composed by steering an H-maser frequency in reference to an ensemble time scale of 18 commercial Cs clocks. Since the ^{87}Sr lattice clock has an accuracy of $<1 \times 10^{-16}$, the discrepancy of green dots and the blue line is attributed to the Cs ensemble. The maximum discrepancy often amounts to 5×10^{-15}, which easily causes the time difference of more than 10 ns in 20 days.

The data shown in Fig. 19.3 enables the simulation of steering the time scale using an optical clock. We assume in the simulation that the HM4 signal is supplied to a phase micro-stepper (PMS), where the frequency error estimated as the red thick line in Fig. 19.3 is compensated; thus, the calibrated output of the PMS is steered by an intermittent operation of an optical clock. A simulation by this method is feasible in our environment because the relative time difference of the HM4 vs UTC(NICT) is monitored and recorded in the JST system on every second, thus enabling the calculation of the time error of the optically steered virtual time scale against UTC. Note that UTC – UTC(NICT) is only available every 5 days. We can compare the optical-based time scale to UTC at 0:00 UTC at every 5 days.

The result of the comparison against UTC is shown in Fig. 19.4. The squares are the real record of UTC(NICT) which are found in Circular T. On the other hand, the circles are the simulation result. The vertical bands represent the duration in which the UTC(NICT) is calibrated by the Sr. We did not make a real physical signal, but the method described above allows to simulate the possible result of generating a time scale. The estimation of the HM4 overall drift rate with reference to the first 4 days was affected by a sporadic short-term frequency fluctuation, which causes the increase in the time difference during MJD 57060–57080. After the second campaign that began on MJD 57079, the overall drift of the HM4 frequency is well estimated, realizing a stable time difference. While the current system of UTC (NICT), which steers the H-maser frequency referring to the Cs ensemble, causes a maximum deviation of 7 ns from UTC during the 65 days of MJD 57074–57139, the steering using an optical clock calibration suppresses the deviation to 3 ns. This improvement is found in the instability as well. The 5-day instabilities of squares and circles after MJD 57079 are 2.7×10^{-15} and 1.3×10^{-15}, respectively.

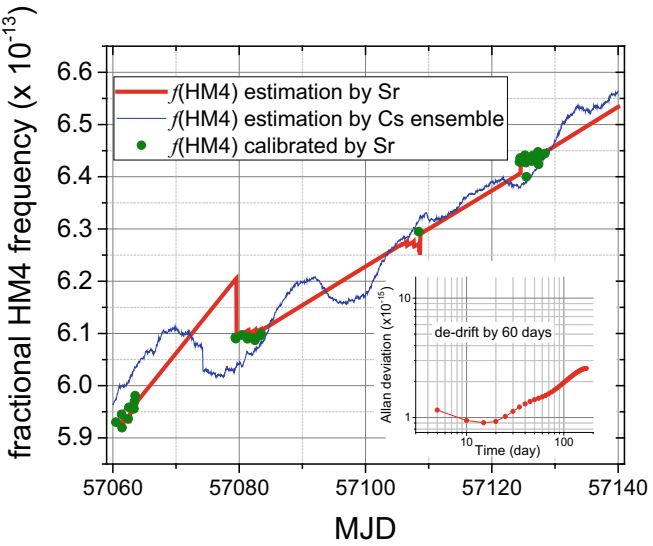


Fig. 19.3 Estimation of a H-maser (HM4) frequency by intermittent operations of an optical clock. The inset shows the instability of the HM4 against UTC, where the drift is removed by the fitting of 60 days beforehand

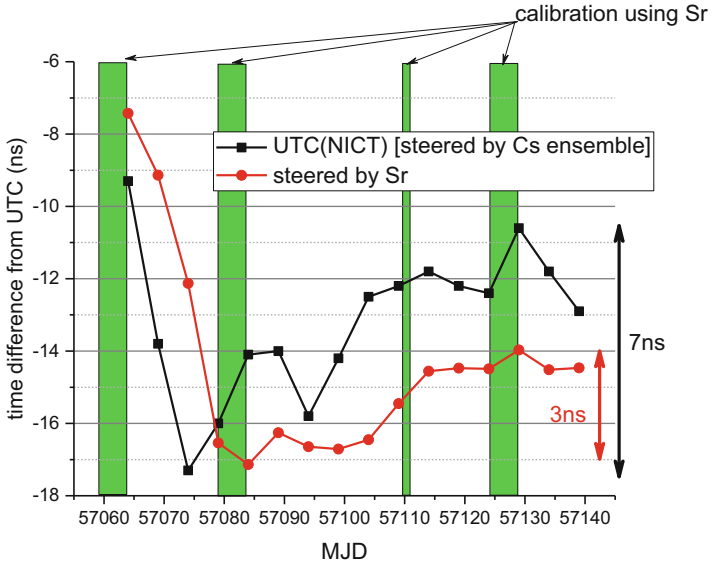


Fig. 19.4 Simulation of a steering of a time scale using an optical clock (*circle*). *Filled squares* are the results reported in Circular T. Note that manual adjustment was made on MJD 57074 for UTC(NICT)

Summary

In summary, we have investigated the feasibility of steering a time scale using an optical clock. We performed three sets of 5 (or 4)-consecutive-day measurements as well as a 1-day measurement of ^{87}Sr lattice clock frequency with reference to an identical H-maser over two and a half months. This is equivalent to evaluating the frequency as well as the drift rate of the H-maser with reference to the optical clock. The optical clock is so stable that 10^4 s of operation is sufficient to evaluate the scale interval of the H-maser at 10^{-15} level. Furthermore, the numerical simulation of the steering showed that intermittent operations of optical clocks for 10^4 s once in 2 weeks can maintain the time scale, allowing an optical frequency standard to be utilized for other applications in extra time. The ability to steer the time scale using an optical clock would allow some laboratories to focus on optical atomic standards. This would be the case particularly for some laboratories, where the study of frequency standards is initiated not in microwave but in optical from the first.

Acknowledgments The authors thank H. Ishijima, S. Ito, K. Kido, and M. Mizuno for the technical help in the measurements. The fruitful discussion with T. Gotoh, H. Ito, M. Kumagai and M. Hosokawa was also a great help for this work.

References

M. Abgrall, B. Chupin, L. De Sarlo, J. Guena, P. Laurent, Y. Le Coq, R. Le Targat, J. Lodewyck, M. Lours, P. Rosenbusch, D. Rovera, S. Bize, C. R. Phys. **16**, 461 (2015)
A. Bauch, S. Weyers, D. Piester, E. Staliuniene, W. Yang, Metrologia **49**, 180 (2012)
H. Hachisu, T. Ido, Jpn. J. Appl. Phys. **54**, 112401 (2015)
S. Peil, J. Hanssen, T.B. Swanson, J. Taylor, C.R. Ekstrom, J. Phys. Conf. Ser. **723**, 012004 (2016)
A. Yamaguchi, N. Shiga, S. Nagano, Y. Li, H. Ishijima, H. Hachisu, M. Kumagai, T. Ido, Appl. Phys. Express **5**, 022701 (2012)

Chapter 20
Activities of Time and Frequency Metrology at NICT: Optical and Microwave Frequency Standards and Their Remote Comparisons

T. Ido, M. Fujieda, H. Hachisu, K. Hayasaka, M. Kajita, M. Kumagai, Y. Li, K. Matsubara, S. Nagano, N. Ohtsubo, Y. Hanado, and M. Hosokawa

Abstract The National Institute of Information and Communications Technology (NICT) provides the national frequency standard of Japan as well as the Japan Standard Time (JST). Besides that, atomic frequency standards, including cesium fountains, a calcium ion clock, a strontium (Sr) lattice clock, and an indium ion clock, are also studied. This contribution briefly summarizes the current status of these frequency standards. We also study the technology to make remote comparison of these standards against others developed in external institutes. Sr lattice clock offers a convenient platform for this purpose because nearly ten Sr-based clocks are available worldwide including the one in NICT with the systematic uncertainty of 8.6×10^{-17}. In 2011, a fiber link to the University of Tokyo was established with the baseline of 24 km, whereas that with a long baseline to PTB (Physikalisch-Technische Bundesanstalt, Germany) was realized in 2013 by a two-way satellite link using the carrier phase. The former revealed the gravitational redshift caused by the elevation difference of 56 m with uncertainty of 7 m. On the other hand, the latter for the first time realized direct comparison of two optical clocks in intercontinental scale.

Keywords Japan Standard Time • Primary clocks • Strontium • Lattice clock • Indium • Ion clock

Japan Standard Time

UTC(NICT), the base of the Japan Standard Time, is a realization of an atomic timescale comprising of an ensemble of 18 Cs commercial atomic clocks (Microsemi Corporation "5071A") Hanado et al. 2008). In this ensemble timescale,

T. Ido (✉) • M. Fujieda • H. Hachisu • K. Hayasaka • M. Kajita • M. Kumagai • Y. Li
K. Matsubara • S. Nagano • N. Ohtsubo • Y. Hanado • M. Hosokawa
National Institute of Information and Communications Technology, 4-2-1 Nukui-kitamachi, Koganei, Tokyo 184-8795, Japan
e-mail: ido@nict.go.jp

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_20

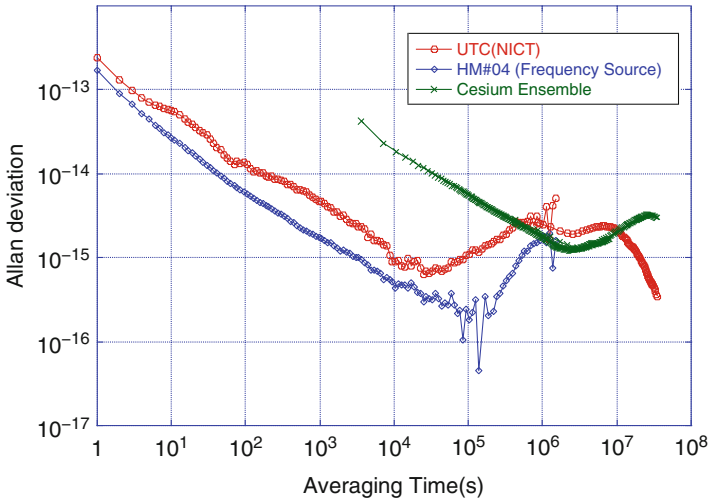


Fig. 20.1 Frequency stability of UTC(NICT) compared with H-maser and Cs ensemble time

the rate of each clock is estimated from the last 30-day trend, and the clock weight is set by $1/\sigma_y$ at $\tau = 10$ days, where σ_y is the Allan deviation. If any clock shows a sudden rate change for more than 1×10^{-14}, its weight becomes zero. For the realization of this Cs ensemble timescale, a phase micro stepper (PMS) stabilized to a hydrogen maser is used. We have four hydrogen masers produced by Anritsu Corporation, and one of them is used as the source of UTC(NICT). The PMS is automatically steered every 8 h to trace the Cs ensemble timescale and is manually steered to trace UTC if necessary. The 5 MHz signals from all clocks in the ensemble are measured using a 24-ch DMTD (multichannel dual-mixer time difference) system with precision of 0.2 ps (Nakagawa et al. 2005). Phase data is measured in addition to the frequency data using one pulse per second (1 PPS) signals to prevent cycle-slip mistakes. For robustness, the main parts of the system have redundancies; atomic clocks and main devices are supplied with a large UPS, a generator which has sufficient fuel to maintain power for 3 days. In addition, the building employs quake-absorbing technologies. Figure 20.1 shows the frequency stability of UTC(NICT) calculated from the data during 2011–2015.

To improve reliability, a distributed generation system of Japan Standard Time is being developed. In this system, atomic clocks at remote stations will be connected via satellites or optical fibers, and an ensemble timescale at each station will be constructed independently from all these connected clocks. As these ensemble timescales at remote stations become approximately the same, they can be used as backup timescales in emergency. This system will ensure a continuity of Japan Standard Time even if the NICT headquarter suffers from a disaster. As the first remote station, JST subsystem that consists of atomic clocks and relevant components has been installed in Kobe branch. Furthermore, time link systems containing

the NICT headquarter, the Kobe branch, and two LF stations have been installed and calibrated, and preliminary operation tests are in progress.

Primary Clocks

NICT has been developing Cs atomic fountain primary frequency standards NICT-CsF1 and NICT-CsF2 for contribution to the determination of TAI and the calibration of Japan Standard Time. While our first fountain CsF1 had been in operation with a typical uncertainly of 1.4×10^{-15} since 2006 (Kumagai et al. 2008), it is currently being upgraded toward an operation at the 10^{-16} level. At first we introduced a cryogenic sapphire oscillator (CSO) developed by the University of Western Australia as an ultra-stable reference. The short-term stability of $6 \times 10^{-14}/\tau^{1/2}$ was realized in high-density operation mode. Additionally, for precise evaluation of the large collisional shift, a rapid adiabatic passage (Dos Santos et al. 2002) was installed in CsF1, enabling both high frequency stability and accuracy. Now we are reevaluating the distributed cavity phase (DCP) shift following the approach proposed in Szymaniec et al. (2007).

In contrast to CsF1 which uses a (0,0,1) laser cooling geometry with quadruple magnetic field, the second fountain CsF2 adopts (1,1,1) geometry enabling many atoms to be captured without a magnetic field gradient in large-diameter laser beams, resulting in a reduction in the atomic density and thus a smaller collisional shift. CsF2 realized a frequency stability of $3 \times 10^{-13}/\tau^{1/2}$ and completed evaluations of most systematic frequency shifts at an uncertainty below 5×10^{-16}. However, the vacuum problem occurred a few years ago. Now the system reconstructions including the vacuum repair are complete and reevaluations of the frequency shift are ongoing. Toward a long-term continuous operation of the fountains for the contribution to a timescale, the CSO was upgraded from dewar type to "cryocooler" type in 2014 (Hartnett and Nand 2010). The cryocooler CSO enables the long-term continuous operation without phase jumps due to liquid helium transfer. To prevent acoustic noise of the pulse tube from disturbing the fountain operations, it is located well away from the fountains, and the only microwave signal is transferred to the fountain room via 100 m optical fiber cable. This ultra-stable signal is converted to 9.192 GHz and used in the microwave interrogation for both fountains.

^{87}Sr Lattice Clock

A lattice clock based on the $^{87}Sr\ ^{1}S_0$-$^{3}P_0$ transition has been in operation since 2011. The absolute frequency of the transition was lately measured to be 429 228 004 229 872.85 (47) Hz (Hachisu and Ido 2015) with reference to the International Atomic Time (TAI). The fractional uncertainty of 1.1×10^{-15} is smaller than any other

TAI-based frequency measurements. The dead time uncertainty of a local transfer oscillator (hydrogen maser) was reduced by homogeneously distributed intermittent measurement over a 5-day grid of TAI. This frequency agrees with previous measurements at other institutes. The Sr optical frequency standard contributes a little with 8.7×10^{-17}, which comprises blackbody shift, lattice Stark shift, dc Stark shift, and density shift. We expect that some straightforward modification of the setup will soon reduce the systematic uncertainty further.

Comparison with UT Using an Optical Fiber Link

Using a telecommunication optical fiber link between NICT and the University of Tokyo (UT) (Fujieda et al. 2011), we performed a direct comparison of two ^{87}Sr clocks (Yamaguchi et al. 2011). Figure 20.2a shows a schematic diagram of the frequency comparison. The clock frequency at NICT is optically transferred to UT using an optical fiber link through the urban region of Tokyo. At NICT, the Ti:sapphire-based optical frequency comb (Ti:Sa comb) is phase-locked to the clock laser at 698 nm. The telecom laser at 1538 nm is phase-locked to the Ti:Sa comb through its frequency-doubled light at 769 nm and transferred to UT through a 60-km-long optical fiber, to which a phase noise cancelation method is applied. The Ti:Sa comb at UT is phase-locked to the frequency-doubled transferred light. The frequency offset between the two Sr clocks is detected at UT as a beat signal between the clock laser and the nearest component of the Ti:Sa comb.

The observed frequency difference is caused by different systematic shifts of the two clocks, in which the differential gravitational redshift of 2.62 Hz, which corresponds to 65 m elevation difference, is the largest. The weighted mean of 11 measurements resulted in the fractional frequency difference between the two distant Sr lattice clocks to be $(1.0 \pm 7.4) \times 10^{-16}$, demonstrating the direct confirmation of the reproducibility of the frequency in distant optical clocks. The uncertainty of the frequency comparison was limited not by the fiber transfer but by

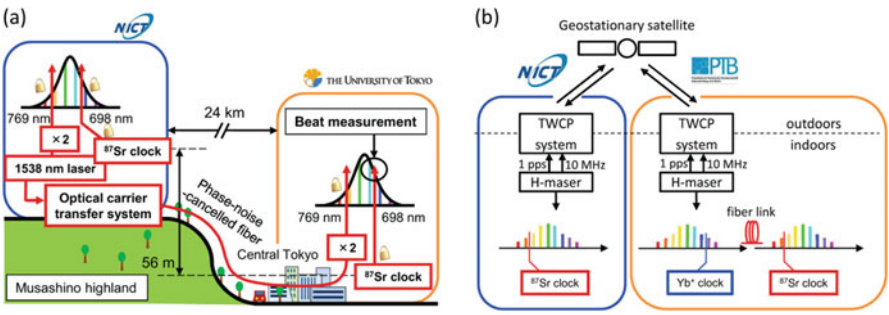


Fig. 20.2 (**a**) Schematic diagram of NICT-UT fiber-link experiment. (**b**) Schematic diagram of NICT-PTB satellite link

the Sr clocks. Details of the optical fiber-link system and the frequency comparison of two distant clocks are described in Fujieda et al. (2011) and Yamaguchi et al. (2011), respectively.

Comparison with PTB Using a TWCP Link

Two ^{87}Sr lattice clocks located at NICT in Japan and at PTB in Germany are directly compared using a satellite-based frequency transfer using the TWCP (two-way carrier phase) (Hachisu et al. 2014). The baseline of 9000 km is the longest ever realized in the direct comparison of optical clocks to the best of our knowledge. Figure 20.2b shows a schematic diagram of the comparison, which involves a link system based on the TWCP technique, optical clocks, and frequency combs at each site. The technical details of the TWCP link are described in Fujieda et al. (2012). The frequency of the lattice clock at each site is counted with reference to a local hydrogen maser (H-maser). A $^{171}Yb^{+}$ clock based on the E3 transition at PTB was employed as a stable transfer oscillator to extend the measurement time when the lattice clock at PTB is offline. A total measurement time resulted in 83,640 s including the extension by the Yb^{+} clock. Both H-masers are linked by the TWCP system. The frequency ratio of the two ^{87}Sr clocks is derived from the ratio of the two local H-masers in real time. We measured the clock frequencies with reference to each local H-maser for several hours per day over 4 days. Taking into account the uncertainty of the ^{87}Sr lattice clock at NICT and PTB, we concluded that a fractional difference of two distant Sr clocks is $(1.1 \pm 1.6) \times 10^{-15}$. Details of the frequency comparison are described in Hachisu et al. (2014).

Ion Clocks

NICT has improved the experimental setup in several ways for the $^{40}Ca^{+}$ ion trap after our first report of the absolute transition frequency in 2008 (Matsubara et al. 2008). Firstly, photo ionization is employed to increase the production efficiency of $^{40}Ca^{+}$ ion trapping. Secondly, a two-layer magnetic shield was installed on the vacuum chamber. Thirdly, mechanical and acousto-optic modulator shutters were installed to ensure thorough reduction in coupling of the cooling laser to the clock transition. Lastly, a fiber noise cancelation technique has been implemented between clock laser and the ion. These improvements resulted in an observed spectral line width of the clock transition of about 30 Hz.

The clock transition frequency was evaluated by microwave link to International Atomic Time (TAI) in more than ten measurement campaigns over a year. The measured frequency of 411 042 129 776 398.4 (1.2) Hz (Matsubara et al. 2012) agrees with the CIPM recommendation (Comité International des Poids et Mesures

2015). Furthermore, using optical comparison against a Sr lattice clock locally available in NICT enabled measurement of the frequency ratio of the $^{40}Ca^{+}$ transition to that of the Sr lattice clock. The result 0.957 631 202 358 049 9(2 3) with fractional uncertainty of 2.4×10^{-15} agrees with the frequency ratio separately evaluated by microwave links to the SI second.

While this frequency was for 5 Hz higher than that measured in Innsbruck (Chwalla et al. 2009) and Wuhan (Huang et al. 2012), the inconsistency was mostly resolved by the latest accurate measurement performed in Wuhan (Huang et al. 2016).

An ion-based clock using $^{115}In^{+}$ is being developed at NICT with an expected inaccuracy in the order of 10^{-18} for the $^{1}S_0$-$^{3}P_0$ transition at 237 nm. A new approach of using two $^{40}Ca^{+}$ ions for sympathetic cooling of the In^{+} as well as for micromotion probe has been employed, enabling the observation of the $^{1}S_0$-$^{3}P_1$ transition for the state detection. The clock transition is lately observed for the first time with the sympathetic cooling scheme, which will be reported elsewhere. A pulsed coherent vacuum ultraviolet (VUV) light source has been developed to excite the $^{1}S_0$-$^{1}P_1$ transition (159 nm) for faster detection of the quantum state (Wakui et al. 2014). This state detection method can extend the single-ion In^{+} clock to a multi-ion clock, which is expected to break through the stability limit of single-ion clocks.

Time and Frequency Transfer

In order to improve the measurement precision of the time and frequency transfer, NICT has studied TWCP (Fujieda et al. 2012) these couple of years. Under cooperation with PTB, the TWCP measurement was performed in the very long baseline of 9000 km (Fujieda et al. 2014). We obtained a short-term instability for frequency transfer of 2×10^{-13} at 1 s, which is at the same level as previously confirmed over a shorter baseline within Japan. Additionally, we confirmed the agreement between TWCP and GPSCP results. A double difference between GPSCP and TWCP shows that the measurement frequency stability reaches the 10^{-16} region. The demonstration of a direct frequency comparison of two Sr lattice clocks was successfully performed by TWCP technique between NICT and PTB (Hachisu et al. 2014). The frequency agreement of the two Sr clocks was confirmed on an intercontinental scale.

Acknowledgments The authors are grateful to the collaboration members of the University of Tokyo and PTB. All experimental works described here were supported by the technical assistance provided by H. Ishijima, S. Ito, and M. Mizuno.

References

M. Chwalla, J. Benhelm, K. Kim, G. Kirchmair, T. Monz, M. Riebe, P. Schindler, A.S. Villar, W. Hänsel, C.F. Roos, R. Blatt, M. Abgrall, G. Santarelli, G.D. Rovera, P. Laurent, Phys. Rev. Lett. **102**, 023002 (2009)

Comité International des Poids et Mesures, CIPM Recommendation 2(CI-2015), 2015, Updates to the list of standard frequencies, http://www.bipm.org/jsp/en/CIPMRecommendations.jsp

F. Dos Santos, H. Marion, S. Bize, Y. Sortais, A. Clairon, C. Salomon, Phys. Rev. Lett. **89**, 233004 (2002)

M. Fujieda, M. Kumagai, S. Nagano, A. Yamaguchi, H. Hachisu, T. Ido, Opt. Express **19**, 16498 (2011)

M. Fujieda, T. Gotoh, F. Nakagawa, R. Tabuchi, M. Aida, J. Amagai, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **59**, 2625 (2012)

M. Fujieda, D. Piester, T. Gotoh, J. Becker, M. Aida, A. Bauch, Metrologia **51**, 253 (2014)

H. Hachisu, T. Ido, Jpn. J. Appl. Phys. **54**, 112401 (2015)

H. Hachisu, M. Fujieda, S. Nagano, T. Gotoh, A. Nogami, T. Ido, S. Falke, N. Huntemann, C. Grebing, B. Lipphardt, C. Lisdat, D. Piester, Opt. Lett. **39**, 4072 (2014)

Y. Hanado, K. Imamura, N. Kotake, F. Nakagawa, Y. Shimizu, R. Tabuchi, Y. Takahashi, M. Hosokawa, T. Morikawa, Int. J. Navig. Obs. **2008**, 1–7 (2008)

J.G. Hartnett, N.R. Nand, IEEE Trans. Microwave Theory Tech. **58**, 3580 (2010)

Y. Huang, J. Cao, P. Liu, K. Liang, B. Ou, H. Guan, X. Huang, T. Li, K. Gao, Phys. Rev. A **85**, 030503 (2012)

Y. Huang, H. Guan, P. Liu, W. Bian, L. Ma, K. Liang, T. Li, K. Gao, Phys. Rev. Lett. **116**, 013001 (2016)

M. Kumagai, H. Ito, M. Kajita, M. Hosokawa, Evaluation of caesium atomic fountain NICT-CsF1. Metrologia **45**, 139 (2008)

K. Matsubara, K. Hayasaka, Y. Li, H. Ito, S. Nagano, M. Kajita, M. Hosokawa, Appl. Phys. Express **1**, 067011 (2008)

K. Matsubara, H. Hachisu, Y. Li, S. Nagano, C. Locke, A. Nogami, M. Kajita, K. Hayasaka, T. Ido, M. Hosokawa, Opt. Express **20**, 22034 (2012)

F. Nakagawa, M. Imae, Y. Hanado, M. Aida, IEEE Trans. Instrum. Meas. **54**, 829–832 (2005)

K. Szymaniec, W. Chałupczak, E. Tiesinga, C.J. Williams, S. Weyers, R. Wynands, Phys. Rev. Lett. **98**, 153002 (2007)

K. Wakui, K. Hayasaka, T. Ido, Appl. Phys. B **117**, 957 (2014)

A. Yamaguchi, M. Fujieda, M. Kumagai, H. Hachisu, S. Nagano, Y. Li, T. Ido, T. Takano, M. Takamoto, H. Katori, Appl. Phys. Express **4**, 802203 (2011)

Chapter 21
IAU Standards of Fundamental Astronomy (SOFA): Time and Date

Catherine Hohenkerk

Abstract The International Astronomical Union's (IAU) SOFA collection of ANSI C and Fortran software provides users with a library of routines that implement official International Astronomical Union (IAU) algorithms for fundamental astronomy computations. These include routines for Earth rotation, precession, nutation, polar motion, coordinate conversions, and astrometry, to list a few. The routines form the various parts of the transformation between the International Terrestrial Reference System (ITRS) and the Celestial Reference System, either the Geocentric (GCRS), Barycentric (BCRS), or the International (ICRS) Celestial Reference System. Most require input parameters that include an "instant"—a time and date in the appropriate time scale. This poster highlights the routines in the categories Time Scales and Calendars, which transform between the time scales UTC, UT1, TAI, TT, TDB, TCB, and TCG as well as provide civil and Julian date conversions.

Keyword Time scales

The SOFA Service

This report first describes the current state of the SOFA service and then highlights the time scales and calendar dates group of routines. Finally, some usage statistics are presented.

The IAU's SOFA service has the task of establishing and maintaining an accessible and authoritative set of algorithms and procedures that implement standard models used in fundamental astronomy. The service is managed by an international panel, the SOFA Board, which is an IAU Functional Working Group appointed through Division A, Fundamental Astronomy. SOFA also works closely with the International Earth Rotation and Reference Systems Service (IERS).

C. Hohenkerk
IAU SOFA Board, HM Nautical Almanac Office, UK Hydrographic Office, Taunton TA1 2DN, UK
e-mail: catherine.hohenkerk@gmail.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_21

SOFA was set up in the 1990s by IAU Division I, the predecessor to Division A. Since the first release in 2001, which contained only Fortran routines, there have been 12 releases. Since the sixth release at the beginning of 2009, there have been both Fortran and ANSI C libraries. All Fortran routines are named iau_ABC, where the letters or numbers ABC is a short string giving some indication of the purpose of the routine and in many cases an epoch. ANSI C routines are named iauAbc. When writing about the routines, it is often just the Fortran form of the name, ignoring the iau_, that is used: ABC. If the name ends in a two-digit number, then the routine is connected to some IAU resolution in a given year; for example, NUT06A refers to the IAU 2006 resolutions. Each routine contains detailed documentation on the arguments, the units, and how the routine should be used, together with relevant references. An alternative to this expert-orientated documentation is the set of Cookbooks, for both Fortran and ANSI C users. There are at present three:

- Tools for Earth Attitude
- Time Scale and Calendar Tools
- Astrometry Tools

They give an overview, which will be particularly helpful to the nonexpert and include explanations, code examples, and test results, to help users understand and check their own application code.

The 12th Release

This latest release, issued on May 3, 2012, consists of some 231 astronomy routines and 55 vector/matrix and utility routines, organized in some 11 categories, which are:

- *Calendars and Time Scales* (27 routines): Conversions between Gregorian and Julian dates and epochs and routines relating the time scales UTC, TAI, UT1, TT, TDB, TCB, and TCG. Further information on these groups is given below.
- *Earth Rotation and Sidereal Time* (15 routines): Routines for ERA, mean and apparent sidereal time and equation of the origins.
- *Precession/Nutation/Polar Motion* (64 routines): IAU 2006/2000A theories; the matrices **C**, **B**, **P,** and **N** (in various combinations); the CIO X, Y, and s; and all the traditional angles. Long-term precession routines ($\pm$200,000 years) of Vondrak, Capitaine, and Wallace were added in this latest release.
- *Geocentric/Geodetic Transformations* (5 routines): Transformations between the WGS 72 and 84, GRS80, and users' own reference ellipsoids, plus various routines for conversion between systems.
- *Ephemerides* (2 routines): Low precision position and velocity of the Earth/Sun, planets, and Earth/Moon barycenter.

- *Astrometry* (36 routines): Space motion, parallax, light deflection, aberration, and the various routines required for the conversion between the various systems, viz., ICRS, BCRS, GCRS, CIRS, TIRS, and ITRS.
- *Galactic Coordinates* (2 routines): Conversions between IAU 1958 and ICRS.
- *Ecliptic Coordinates* (6 routines): Conversion between equatorial and ecliptic coordinates. This is a new group in the twelfth release.
- *Fundamental Arguments* (14 routines): These are the 14 fundamental arguments, i.e., the mean longitude Sun, Moon, planets, elongation Sun minus Moon, etc. used in the nutation series and given by the IERS's 2003 conventions.
- *Star Catalog Conversions* (5 routines): Conversions between FK5 and Hipparcos catalog positions.
- *Vector/matrix/utility* routines (55 routines):

Those routines that implement IAU Standards are designated as **canonical** (there are 59 in this category), which is indicated by a two-digit year in the name, e.g., ERA00 and PMN06A.

The Time Scale and Calendar Routines

SOFA has several routines for converting between the Gregorian calendar dates and Julian dates (JD). Julian dates, as opposed to Julian day numbers, include the time as a fractional part. The SOFA code always expresses Julian dates (or quasi-Julian dates when dealing with UTC and leap seconds) as two double precision numbers, the sum of which is the desired JD, including the fraction. These are handled in ways that preserve precision, important when dealing with microseconds. Other calendar routines deal with Julian and Besselian epochs.

The time scales that SOFA recognizes are listed in Table 21.1. SOFA chose the simple scheme illustrated in Fig. 21.1, not only because each of these seven time scales has a distinctive nature and use but also to give the user the maximum flexibility and control while maintaining correctness.

It is worth noting that these routines include the transformations to and from UTC and consequently have to deal with leap seconds. A positive leap second is formatted as 23:59:60.0000 (with as many decimals as is necessary for the user application and within the limit of the computer architecture).

There are two important routines that are part of these time scale conversions: DAT and DTDB. DAT—a canonical routine—deals with the difference between UTC and TAI. It must be updated whenever a leap second is added or removed. It is automatically called whenever a conversion is made between UTC and TAI. Routine DTDB, which delivers TDB–TT, is not automatically called. Users must supply TDB–TT as an argument to the routine converting between TT and TDB and thus may either use the model provided by SOFA or one of their own choosing. Time scale support was first included in the 2010 December release.

Table 21.1 Summary of time scales

	Names of time scales	Nature
UT1	Universal time based	Earth rotation
UTC	Coordinated universal time	Atomic/Earth hybrid
TAI	International atomic time	Atomic
TT	Terrestrial time	Dynamical
TCG	Geocentric coordinate time	Dynamical
TCB	Barycentric coordinate time	Dynamical
TDB	Barycentric dynamical time	Dynamical

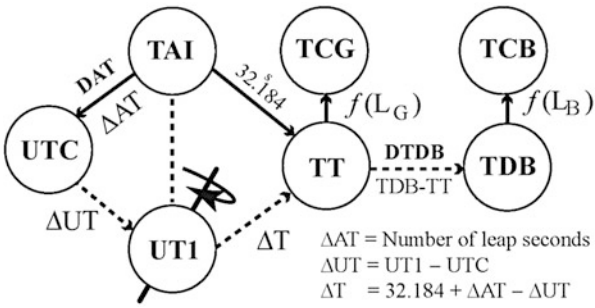


Fig. 21.1 Linkage between time scales. The *dotted lines* indicate the relationships that depend on the user providing the appropriate linking value. UT1–UTC is provided by the IERS. The functions f(LG) and f (LB) represent IAU defined linear models

The Use of SOFA

There are a few groups that use other implementations of SOFA. It should be stressed that under SOFA's terms and conditions users may change any of the code, but if they do, then at least they must remove the letters "iau" from the name of the routine so that the modified form cannot be inadvertently linked to an application that is expecting the standard form. This strategy is vitally important since SOFA code, as downloaded, forms an IAU standard, and any change of code is then not part of that IAU standard.

There are three implementations listed on the SOFA website. One is a JAVA version by Paul Harrison, from the Jodrell Bank Centre for Astrophysics at the University of Manchester. There is a C# version for the World Wide Astronomy (WWA) library, implemented by Attila Abrudan. Lastly, there is the Essential Routines for Fundamental Astronomy (ERFA) implemented by the AstroPy group. This version has modified routine names in order to be compatible with the typical free/open-source software licenses that unfortunately do not recognize the concept of "read-only software."

Currently SOFA uses the AWStats web tool to analyze usage. However, it must be borne in mind that website statistics can be misleading. It is not necessarily a record of individuals use since it is unclear whether it is an individual downloading one of the libraries or a system manager setting up the SOFA library for a group. Having said that, the use of the SOFA website has increased from an average of 769 unique visitors per month in 2008 to over 1900 in 2015. There have been over

15,000 downloads of the previous release, 11a (both Fortran and ANSI C), over 14 months. There are currently 760 registered users. All users are encouraged to register, as this allows them to be notified when there are issues and releases.

The SOFA libraries provide the astronomical community with a set of well-tested independent standard routines that support IAU resolutions. This gives users the tools for easily implementing fundamental astronomy transformations, such as Earth attitude, for use in research, in applications, and importantly in testing. We ask users to acknowledge their use of SOFA in their publications.

Acknowledgments We thank and acknowledge the work of the Board members and their institutes. We are also grateful to the UK Hydrographic Office for hosting the SOFA website. The SOFA Board members are:

John Bangert, US Naval Observatory (retired)
Steven Bell, HM Nautical Almanac Office, UKHO
Nicole Capitaine, Paris Observatory
William Folkner, Jet Propulsion Laboratory
Mickaël Gastineau, Paris Observatory, IMCCE
Catherine Hohenkerk, HM Nautical Almanac Office (Chair)
Jinling Li, Shanghai Astronomical Observatory
Brian Luzum, US Naval Observatory (IERS)
Zinovy Malkin, Pulkovo Observatory, St. Petersburg
Jeffrey Percival, University of Wisconsin
Wendy Puatua, US Naval Observatory
Scott Ransom, National Radio Astronomy Observatory
Patrick Wallace, RAL Space (retired)

Chapter 22
Earth's Variable Clock

L.V. Morrison, F.R. Stephenson, and C. Hohenkerk

Abstract Ancient Babylonian clay tablets buried for centuries beneath the sands of the desert are part of an extensive historical archive that contains vital information about the Earth's rotation from 720 BC to the present. These historical observations of solar and lunar eclipses and occultations of stars are reanalyzed to determine the error in the Earth's clock by the parameter ΔT.

Keywords Earth rotation • Length of day • Delta T

Summary

The poster showed various diagrams displaying the individual results for ΔT, together with the best-fitting curve through these data, leading to the derivation of the changes in the length of day (lod) 720 BC to the present. This analysis is part of a paper that has been submitted to the Proc. R.S. (Lond.). The results, tables of ΔT, length of day, polynomial coefficients, etc. are available from H.M. Nautical Almanac's website at astro.ukho.gov.uk/nao/lvm.

Acknowledgments We are grateful to all those who over the centuries have made and recorded these observations, to Herald and Gault for their modern reanalysis of the data held at the Centre de Données Astronomiques de Strasbourg, and also to DID at the UK Hydrographic Office for printing the poster.

L.V. Morrison • F.R. Stephenson • C. Hohenkerk (✉)
Pevensey, East Sussex, UK

Durham University, Durham, UK

HM Nautical Almanac Office, UK Hydrographic Office, Taunton, UK
e-mail: catherine.hohenkerk@ukho.gov.uk

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_22

Chapter 23
The Determination of Earth Orientation by VLBI and GNSS: Principles and Results

Nicole Capitaine

Abstract The Earth Orientation Parameters (EOP) connect the International Terrestrial Reference System (ITRS) to the Geocentric Celestial Reference System (GCRS). These parameters, i.e., Universal Time, UT1, and pole coordinates in the ITRS and in the GCRS, describe the irregularities of the Earth's rotation. They are mainly determined by two modern astro-geodetic techniques, VLBI (Very Long Baseline Radio Interferometry) on extragalactic radio sources, which is used to realize and maintain the International Celestial Reference System (ICRS), and Global Navigation Satellite System (GNSS), especially GPS (Global Positioning System), which has an important contribution to the realization of the ITRS. The aim of this presentation is twofold: to present the modern bases for the consideration of Earth orientation and to discuss how the principles of VLBI and GPS give access to the measure of different components of the EOP variations, especially UT1. The accuracy that can be achieved is based on the improved concepts, definitions, and models that have been adopted by IAU/IUGG resolutions on reference systems and Earth's rotation, as well as on the refined strategy of the observations.

Keywords Earth rotation • Reference systems • Time • UT1 • IERS • VLBI • IVS • GNSS • IGS

Introduction

The Earth's orientation in space reflects the orientation of the terrestrial reference system relative to a geocentric celestial reference system. Accurate knowledge of that orientation is essential for space mission launches, the reduction of astro-geodetic observations, and scientific exploitation. Determining and providing that orientation is currently coordinated at an international level by the International Earth Rotation and Reference Systems Service (IERS). The IERS products are

N. Capitaine (✉)
SYRTE, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC, 61, avenue de l'Observatoire, 75014 Paris, France
e-mail: nicole.capitaine@obspm.fr

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_23

based on data provided by the international space geodetic services (IVS, ILRS, IGS, IDS), which are derived from observations by various modern techniques, namely, Very Long Baseline Interferometry (VLBI) on extragalactic radio sources for IVS, laser ranging on artificial satellites and the Moon for ILRS, observations with the Global Navigation Satellite System (GNSS) for IGS, and observations with the DORIS system for IDS. In contrast, optical astrometry on which the determination of Earth's rotation has been entirely based until the 1970s is no longer used for that purpose due to its poor accuracy as compared to modern techniques.

Section "Modern Bases for the Consideration of the Earth Orientation" presents the modern bases for the consideration of the Earth orientation as adopted by recent IAU and IUGG resolutions. Section "Earth Rotation Variations and Observations" reports on the Earth rotation variations and observations. Section "Principles of the VLBI and GNSS Techniques for Measuring Earth Rotation" describes the principles of the VLBI and GNSS observations and discusses how they give access to the determination of the different components of Earth's rotation, especially UT1. Section "Comparisons and Results" reports on the accuracy that can be achieved by these techniques, while Section "Summary" gives a summary.

Modern Bases for the Consideration of the Earth Orientation

Earth's Orientation and Reference Systems

The Earth's orientation in space is traditionally represented by five *Earth Orientation Parameters* (EOP), which provide the direction of the pole in the terrestrial reference system due to polar motion, the direction of the pole in the celestial reference system due to precession-nutation, and the variations in the Earth's diurnal rotation based on Universal Time, UT1. The definition and realization of the astronomical reference systems and time scales and the way the Earth's orientation/rotation is expressed and the models that are adopted for the parameters are crucial for any consideration of the Earth's rotation.

The Recent IAU/IUGG Resolutions on Reference Systems and Earth's Rotation

In 1997, the IAU adopted (IAU 1997, 1999 Resolution B2, cf. IAU 1999) the *International Celestial Reference System* (ICRS) based on VLBI observations of extragalactic radio sources and the corresponding frame, the *International Celestial Reference Frame* (ICRF, Ma et al. 1998), that realizes the ICRS.

IAU 2000/IUGG 2003 Resolutions	IAU 2006 & IUGG 2007 Resolutions	IAU 2009 Resolutions	Aim (for consistency with μas accuracy)
Resolution B1.3 *Definition of BCRS and GCRS* Resolution B1.5 *Extended Relativistic framework for time transformation*	Resolution 2: *Definition of GTRS (and ITRS* (International Terrestrial Reference system) as a specific GTRS	Resolution B2 *Adoption of ICRF2*	*Improvement in the definition of the astronomical reference systems*
Resolution B1.6 *IAU 2000 Precession-Nutation Model*	Resolution B1 *Adoption of the P03 Precession and definition of the ecliptic*	Resolution B1 *Adoption of the IAU 2009 System of astronomical constants*	*Adoption of high accuracy astronomical models*
Resolution B1.7 *Definition of Celestial Intermediate Pole* Resolution B1.8 *Definition and use of CEO and TEO*	Resolution B2 *Supplement to the IAU 2000 Resolutions on reference systems* Rec 1: *Harmonizing «intermediate » to the pole and the origin* Rec 2: Default orientation of the BCRS/GCRS		*Refinement in the concepts and definition of the EOP*
Resolution B1.9 *Re-definition of Terretrial Time (TT)*	Resolution B3 *Re-definition of Dynamical Barycentric Time (TDB)*		*Improvement in the defintion of coordinate time scales*

Fig. 23.1 The IAU 2000–2009 and IUGG 2003–2007 resolutions on reference systems

Then, from 2000 to 2009, IAU and IUGG resolutions have been passed (see Fig. 23.1), the purpose of which was to comply with the accuracy and improved properties of the ICRS as well as with the precision of modern astro-geodetic observations. These resolutions have improved the definition of the reference systems and time scales for astronomy, have modified the way the Earth's orientation is expressed, and have adopted high-accuracy models for expressing the relevant quantities for the transformation from terrestrial to celestial systems. This formed the modern bases for the consideration of Earth rotation.

According to these IAU/IUGG resolutions (for more details see IAU 2000, 2001; IUGG 2003; IAU 2006, 2008; IUGG 2007; IAU 2009, 2010; IUGG 2011):

- The systems of space-time coordinates for the solar system and the Earth within the framework of general relativity are named the *Barycentric Celestial Reference System* (BCRS) and the *Geocentric Celestial Reference System* (GCRS), respectively (IAU 2000 Resolution B1.3).
- *Barycentric Coordinate Time* (TCB) and *Geocentric Coordinate Time* (TCG) are the time coordinates of the BCRS and GCRS, respectively (IAU 2000 Resolution B1.3).
- The BCRS-to-GCRS transformation is specified as an extension of the Lorentz transformation for the space and time coordinates (IAU 2000 Resolution B1.3; for more details, see Soffel et al. 2003).
- For all practical applications, unless otherwise stated, the BCRS is assumed to be oriented according to the ICRS axes, and the orientation of the GCRS is derived from the ICRS-oriented BCRS (IAU 2006 Resolution B2).
- *Terrestrial Time* (TT) and *Barycentric Dynamical Time* (TDB) are defined by a linear conventional relationship to TCG (IAU 2000 Resolution B1.9) and TCB (IAU 2006 Resolution B3), respectively.
- The preferred 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 EOP is the *International Terrestrial Reference System* (ITRS, IUGG 2007 Resolution 2).
- The Second Realization of the International Celestial Reference Frame (ICRF2) replaced the ICRF (IAU 2009 Resolution B3) as the fundamental astrometric realization of the ICRS on 1 January 2010; it was endorsed by IUGG 2011 Resolution 3.
- The pole of the nominal rotation axis is the *Celestial Intermediate Pole* (CIP); it is defined as being the intermediate pole in the transformation from the GCRS to the ITRS (IAU 2000 Resolution B1.7).
- The CIP is defined (IAU 2000 Resolution B1.7) as being the intermediate pole, in the GCRS-to-ITRS transformation, separating nutation from polar motion by a specific convention in the frequency domain, such that (1) the celestial motion of the CIP (i.e., precession/nutation) includes all the terms with periods greater than 2 days in the GCRS and (2) its terrestrial motion (i.e., polar motion) includes all the terms with periods less than 2 days in the GCRS.
- Nutation is taken from the MHB2000 nutation model (Mathews et al. 2002) to become IAU 2000A nutation (IAU 2000 Resolution B1.6), which came into force on 1 January 2003.
- Precession is taken from the P03 precession model (Capitaine et al. 2003) as the replacement for the precession part of IAU 2000A, to become IAU 2006 precession (IAU 2006 Resolution 1), which came into force on 1 January 2009.
- The *Earth rotation angle* (ERA) is defined as the angle measured along the equator of the CIP between the specific zero points of celestial and terrestrial longitude (i.e., right ascension, IAU 2000 Resolution B1.8), named the *Celestial Intermediate Origin* (CIO) and the *Terrestrial Intermediate Origin* (TIO), respectively, by IAU 2006 Resolution B2.

- The transformation between the ITRS and GCRS is specified by the position of the CIP in the GCRS, the position of the CIP in the ITRS, and the ERA; this is referred to as *the CIO-based paradigm*.
- UT1 is defined by a conventional relationship as being linearly proportional to the ERA.
- The GCRS position of the CIP can be provided by expressions as a function of time of the rectangular coordinates, X and Y, of the GCRS direction of the CIP unit vector, which include precession and nutation and the frame bias between the pole of the GCRS and the CIP at J2000.0.

The models, procedures, data, and software to implement these resolutions operationally had been made available by the IERS (IERS Conventions 2003 and then 2010) and the Standards of Fundamental Astronomy (SOFA) activity (Wallace 1998). The new terminology associated with these resolutions was recommended by the IAU Working Group on "Nomenclature for Fundamental Astronomy" (Capitaine et al. 2007). This provides a new framework for dealing with Earth's rotation or computing directions of celestial objects or Earth's satellites in various systems, which has to be taken into account in the estimation of the EOP by any of the IERS techniques.

The Fundamental Celestial and Terrestrial Reference Systems

The ICRF, regarded as the realization of the ICRS, is based upon an ensemble of very distant extragalactic radio sources, which do not exhibit any proper motion detectable with the current precision of VLBI observations. It was aligned with the previous IAU official celestial reference system, the FK5 at J2000, but no attempt was made to refer the positions of the sources to the mean pole or mean equinox at a fundamental epoch as was the case in the fundamental catalogs. As a consequence, the celestial reference frame is no longer dependent on the Earth's motion. The ICRF2 (Fey et al. 2015) contains 3414 radio sources of which 295 are *defining sources* (see Fig. 23.2); it represents a significant improvement over the previous version while preserving, by construction, the same directions of the reference axes. The ICRF2 resulted from 6 000 000 VLBI observations performed from 1979 to 2009; its accuracy, about 40 μas, is 5–6 times better than that of the ICRF. Its axis stability is 10 μas, which is nearly twice as stable as ICRF(1), which was itself 100 times as stable as the FK5.

The ITRF, regarded as the realization of the International Terrestrial Reference System, is based on time series of station positions and EOP provided by the IERS Technique centers for the four space geodetic techniques (VLBI, SLR, GNSS, and DORIS) (see Fig. 23.3). The latest version, ITRF2014 (Altamimi et al. 2016), which is based on completely reprocessed solutions of those techniques, is an improved solution as compared to the previous version, ITRF2008, while preserving, by construction, the same directions of the reference axes.

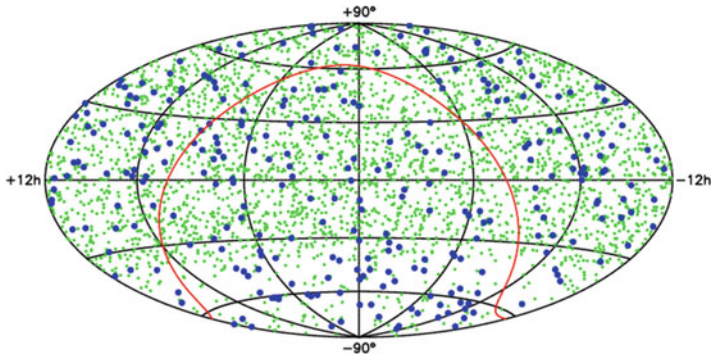


Fig. 23.2 The distribution of the ICRF2 sources on the celestial sphere (Fey et al. 2015). The 295 defining sources are shown in *blue* (source: International Earth Rotation and Reference Systems Service, IERS)

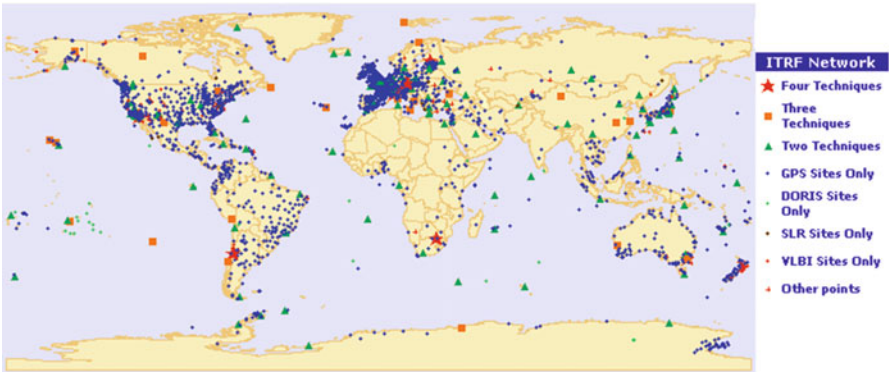


Fig. 23.3 Geographical distribution of the ITRF2008 sites (source: ITRF website)

The Relativistic Space-Time Reference Systems

The BCRS should be used, with TCB, as a global coordinate system for the solar system. In contrast, the GCRS should be used, with TCG, as a local coordinate system for the Earth, e.g., for expressing Earth's rotation or the motion of an artificial satellites around the Earth. The GCRS is defined such that the transformation between BCRS and GCRS spatial coordinates contains no rotation component. The GCRS is therefore, "kinematically non-rotating" with respect to the BCRS so that the equations of motion with respect to the GCRS will contain relativistic Coriolis forces that come mainly from geodesic precession.

The TCB and TCG are the time scales recommended for the solar system and the Earth; however, TT and TDB may be more convenient to use for some practical applications. This is the case, e.g., for the solutions of the Earth's rotational equations that are usually expressed in TT and the solar system ephemerides

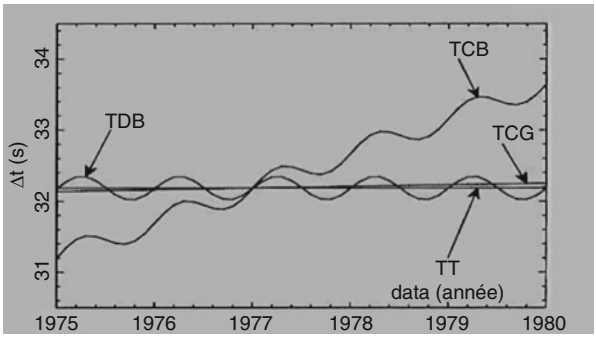


Fig. 23.4 Differences of TT (horizontal line), TCG, and TDB to TAI, in s (not to scale)

(necessary for computing the lunisolar and planetary torque acting on Earth's rotation) that are often expressed in TDB. The TT and TDB can be used with the same rigorous approach as TT and TCB thanks to the IAU-adopted conventional relationship definitions:

$$\mathrm{TCG} - \mathrm{TT} = \mathrm{LG} \times (\mathrm{JD} - 2\,443\,144.50) \times 86\,400\ \mathrm{s}, \tag{23.1}$$

where $L_G = 6.969290134 \times 10^{-10}$ is a defining constant.

$$\mathrm{TCB} - \mathrm{TDB} = \mathrm{LB} \times (\mathrm{JD} - 2\,443\,144.500\,372\,5) \times 86\,400\ \mathrm{s} + \mathrm{TDB}_0 \tag{23.2}$$

where $L_B = 1.550519768 \times 10^{-8}$ and $TDB_0 = 6.55 \times 10^{-5}$ s are defining constants.

Note that (1) the use of the SI second for TT as well as TDB has been recommended by Klioner et al. (2010) and (2) after 1 cy: TCG ~ TT + 2 s; TCB ~ TDB + 50 s (see Fig. 23.4).

The Modern Definition of the EOP and the Coordinate Transformation Between ITRS and GCRS

According to the CIO-based paradigm, adopted by the IAU 2000 Resolutions, the transformation between the ITRS and GCRS is specified by the newly defined EOP, i.e., the position of the CIP in the GCRS and in the ITRS (see Fig. 23.5), and the Earth rotation angle (ERA, see Fig. 23.6):

– The position of the CIP in the GCRS is represented by the x and y coordinates of the CIP unit vector in the GCRS, denoted X and Y; these quantities are often multiplied by $1\ 296\ 000''/2\pi$ in order to represent the approximate values in arcseconds of the corresponding angles with respect to the GCRS z-axis. This can also be expressed in the form of polar coordinates d and E such that $X = \sin d \cos E$ and $Y = \sin d \sin E$ (see Fig. 23.5).

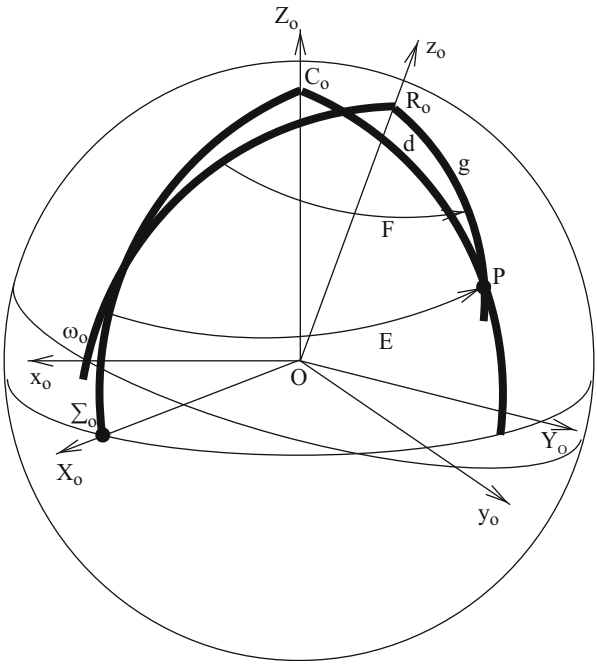


Fig. 23.5 The coordinates of the Celestial Intermediate Pole (CIP) in the GCRS and the ITRS. The CIP is represented by *P*

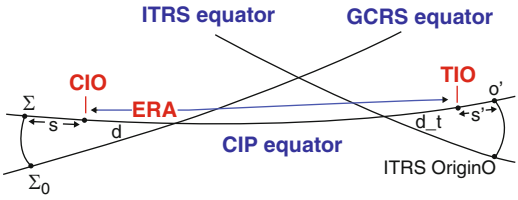


Fig. 23.6 The Earth rotation angle (ERA) along the CIP equator

– The position of the CIP in the ITRS is represented by the x and y coordinates of the CIP unit vector in the ITRS, denoted x and y, in arcseconds, the values of which represent the corresponding angles with respect to the ITRS z-axis. 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. This can also be expressed in the form of the polar coordinates g and F, such that $x = \sin g \cos F$ and $y = -\sin g \sin F$, which are the terrestrial counterpart of the d and E angles in the GCRS (see Fig. 23.5).
– The ERA, which links the two systems, is expressed through the following conventional linear transformation of UT1 (Capitaine et al. 2000):

$$\begin{aligned}&\mathrm{ERA(UT1)}\\&= 2\pi\,[0.7790572732640 + 1.00273781191135448 \times (\text{Julian UT1date} - 2\,451\,545.0)].\end{aligned} \tag{23.3}$$

This linear relationship is a consequence of the kinematically non-rotating nature of the origins (Guinot 1979; Capitaine et al. 1986) to which the ERA refers. The positions of the CIO and the TIO are provided by the *CIO and TIO locators* that are designated s and s' (see Fig. 23.6).

The CIO-based coordinate transformation to be used to transform from the ITRS to the GCRS at epoch t of the observation can be written as the following matrix product:

$$[\text{GCRS}] = Q(t) \cdot R(t) \cdot W(t) \cdot [\text{ITRS}], \tag{23.4}$$

where $Q(t)$, $R(t)$, and $W(t)$ are the transformation matrices arising from the motion of the CIP in the celestial reference system and the CIO locator (i.e., X, Y, s), from the Earth rotation angle around the CIP axis (i.e., ERA), and from polar motion and the TIO locator (i.e., x, y, s'), respectively; these can be written as:

$$\begin{aligned} Q(t) &= Q(X, Y, s) = R_3(-E) \cdot R_2(-d) \cdot R_3(E) \cdot R_3(s); \\ W(t) &= R_3(-s') \cdot R_2(x) \cdot R_1(y); R(t) = R_3(-\text{ERA}). \end{aligned} \tag{23.5}$$

The Modern Definition of UT1 and Its Relation with UTC and TAI

According to the new definition of the EOP adopted by the IAU 2000 Resolutions, UT1 is an angle. It is the angle of the Earth's rotation about the CIP axis defined by its conventional linear relation (23.3) to the ERA; that relationship was selected so that UT1 follows in the long term the Greenwich mean solar time.

The Universal Time, UT1, can also be regarded as a time determined by the rotation of the Earth. It can be obtained from the uniform time scale UTC (Coordinated Universal Time), which is related to the International Atomic Time (TAI), by using the quantity UT1–UTC, estimated by observations, that is provided by the IERS.

Note that (1) UT1, UTC, and TAI were coincident on 1 January 1958, (2) TT = TAI + 32.184 s, (3) TAI − UTC has taken the integer values between 10 s and 36 s during the period 1972-2016, and (4) TGPS − UTC = 17 s, TGPS being the GPS time (see Fig. 23.7).

Earth Rotation Variations and Observations

The rotation of the Earth has many irregularities due to the many perturbations that affect the total angular momentum of the Earth, including the elastic and fluid components. Its diurnal rotation has various fluctuations over time; in addition, its

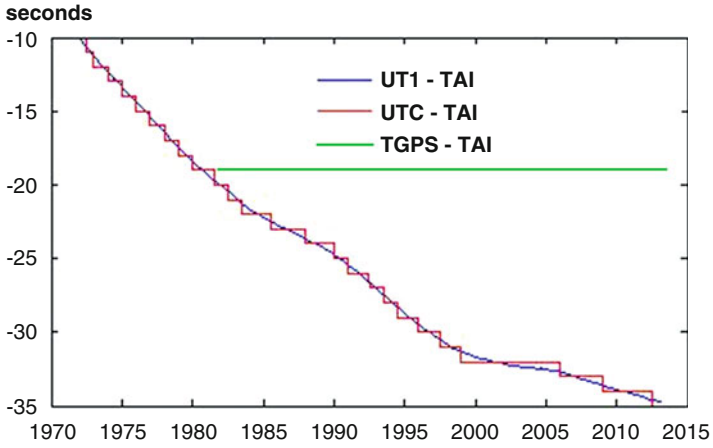


Fig. 23.7 Differences of UT1, UTC, and TGPS to TAI in s (source: Yao 2013)

rotation axis moves in space due to precession-nutation and within the Earth due to polar motion. Important progress in the knowledge of these irregularities has been possible, thanks to the improvement in the precision of the observations, the potential of the various techniques used, and the accuracy of the realization of the space and time reference systems to which the measurements are referred. It has also benefited from the contribution of different branches of geophysics in the measurement and modeling of global geophysical parameters (e.g., atmospheric, oceanic, or hydrostatic angular momentum).

Polar Motion

The main components of polar motion, which have been known for over a century, consist of a quasi-circular free motion with a period of approximately 430 days, called *Chandler wobble*, with amplitudes varying between 30 and 270 mas and an annual elliptical forced motion of amplitude of about 100 mas due to seasonal atmospheric, oceanic, and hydrostatic excitation. The proximity of the frequencies of the Chandlerian and annual terms results in a 6.6-year spiral motion of the pole as seen from the Earth, its position being contained within a square of side 20 m (~0.7″) (see Fig. 23.8). There is also a secular drift of about 4 mas/year toward Canada, which is mainly attributed to the effect of the "postglacial rebound" after the last ice age, about 10 000 years ago. In addition, there are a large number of variations with amplitudes between 1 μas and 20 mas covering the whole range of frequencies corresponding to periods between 0.5 days and several decades. This includes sub-daily variations caused by ocean tides and periodic motions driven by gravitational torques with periods less than 2 days.

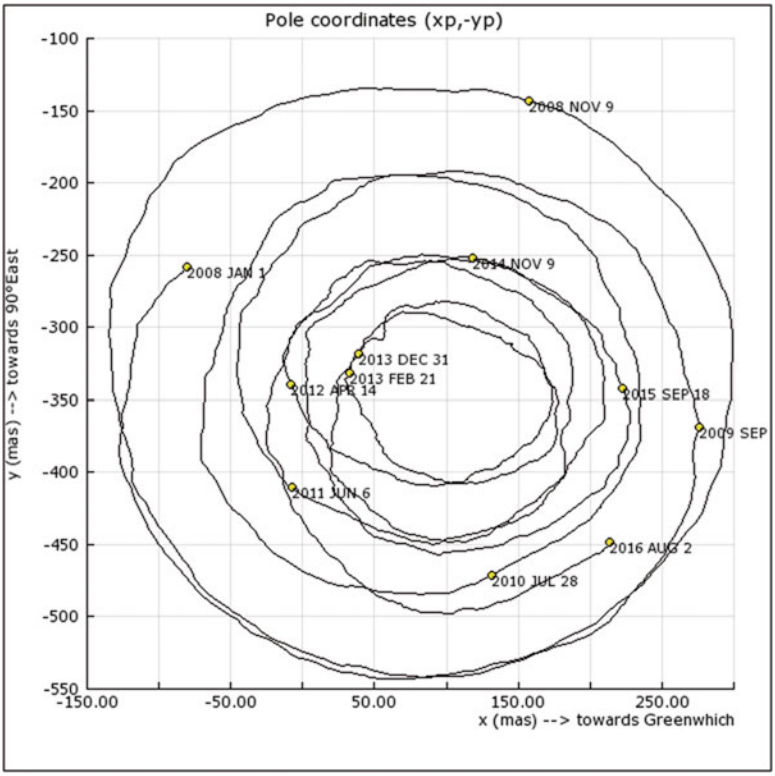


Fig. 23.8 The polhody corresponding to polar motion from January 2008 to August 2016 (source: IERS EOP Product Center, Paris Observatory; http://hpiers.obspm.fr/eop-pc/products)

Precession-Nutation

The motion of the CIP in the GCRS is mainly due to the astronomical phenomenon called precession-nutation (see Fig. 23.9) resulting from the effect of the external torque exerted by the Moon, the Sun, and to a lesser extent the planets, on the equatorial bulge of the Earth. The *precession of the equator*,[1] or equivalently of the CIP, is due to the constant and secular part of the external torque, while the nutation is due to the variable part of that torque. Precession-nutation may be represented with great accuracy by a semi-analytical model in which the precession part is, by convention, the sum of the secular terms, while the nutation includes all the periodic terms. One should also take into account the relativistic rotation of the GCRS—that is, by definition, kinematically non-rotating—with respect to a geocentric celestial

[1]This as to be clearly distinguished from the *precession of the ecliptic*, i.e., the secular (or long term) displacement of the ecliptic which is produced by the planetary perturbations.

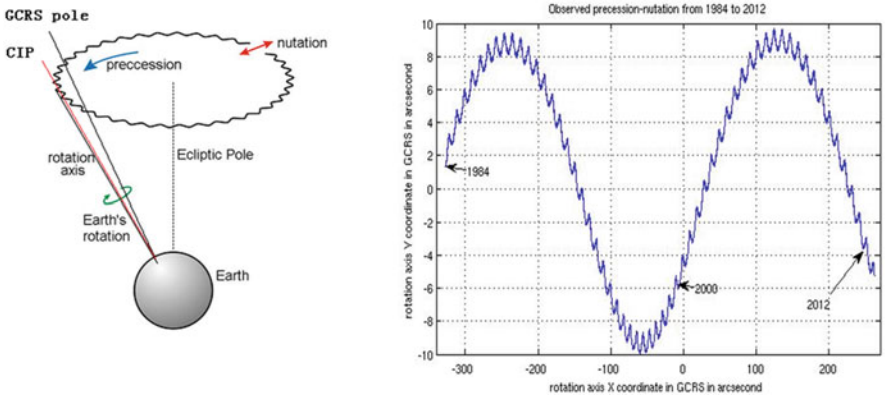


Fig. 23.9 The Earth's precession-nutation: theoretical motion (*left*) and observed motion (*right*) (source: *right*: IERS EOP Product Center, Paris Observatory; http://hpiers.obspm.fr/eop-pc/products)

reference system, that is dynamically non-rotating in the framework of general relativity.

The current IAU precession-nutation model, labeled 2006/2000, is a semi-analytical model, composed of the IAU 2006 precession and the IAU 2000 nutation. The motion of the CIP in the GCRS also includes the celestial term corresponding to the free core nutation (FCN). This free retrograde nutation, with a 430-day period in space, is not predictable and is therefore not included in the IAU model.

The IAU 2000A[2] nutation (Mathews et al. 2002; MHB) consists in a series of 1365 elliptical lunisolar and planetary terms. It contains "in-phase" and "out-of-phase" components, with amplitudes between 0.1 μas and 17.2″, and periods between 3 days and 933 centuries, which can be also be expressed in the form of pairs of prograde and retrograde circular nutations. The periods of the largest three terms are 18.6 years, 0.5 year, and 13.6 days. This model has been obtained by applying the MHB transfer function for nonrigid Earth effects, based on basic Earth parameters (BEP) fitted to VLBI data, to solutions for a rigid Earth obtained by celestial mechanics. The BEP include dynamical ellipticity of the Earth and its fluid core, coefficients of deformability of the whole Earth and its fluid core under tidal forcing, as well as core mantle and outer core to inner core couplings due to the magnetic fields. The nonrigidity effects are between a few μas and a few tens of mas. The resolution and accuracy of the coefficients are 0.1 μas and about 20 μas, respectively. The triaxiality of the Earth is responsible for nutations with a 0.5-day period that are not part of the IAU model but, according to the conventional

[2]The letter A, which will be omitted in the following, is for the full development with an accuracy of 0.1 mas, as opposed to the letter B which is for a reduced development ensuring an accuracy of only 1 mas.

definition of the CIP, are considered through the corresponding terms in polar motion.

The IAU 2006 precession of the equator (Capitaine et al. 2003) was derived from the dynamical equations expressing the motion of the mean pole about the ecliptic pole. The convention for separating precession from nutation and the integration constants used in solving the equations have been chosen in order to be consistent with the IAU 2000 nutation. The IAU 2006 precession includes the contributions to the precession rates from the second order effects, the J_3 and J_4 effects of the lunisolar torque, the J_2 and planetary tilt effects, as well as the tidal effects, the nonlinear terms due to coupling effects, and the geodesic precession. It also includes the Earth's J_2 rate effect ($\mathrm{d}J_2/\mathrm{dt} = -3 \times 10^{-9}$ cy^{-1}), mostly due to the postglacial rebound, which was not taken into account in the previous IAU precession models. The precision of the IAU 2006 precession is expected to be of the order of 0.1 mas over 1 century (Hilton et al. 2006).

It should be clear that this model is intended for high-accuracy applications over a limited time span (i.e., a few centuries around the fundamental epoch). This is due to the fact that, like its predecessors, this IAU model is given as a set of polynomial approximations, while precession in fact is the sum of quasiperiodic motions with very long periods. Precession expressions, compatible with IAU 2006 but intended for applications over long-term intervals, have been developed (Vondrák et al. 2011) to ensure an accuracy of a few arcseconds throughout the historical period and a few tenths of a degree after several thousand centuries.

The IAU 2006/2000 expressions for the X and Y quantities, which include precession, nutation, and frame bias (w.r.t. the GCRS), can be written as follows (Capitaine and Wallace 2006):

$$
\begin{aligned}
X = {} & -0.016617 + 2004.191898t - 0.4297829t^2 \\
& - 0.19861834t^3 - 0.000007578t^4 + 0.0000059285t^5 \\
& + \Sigma_i \left[(a_{s,0})_i \sin(\text{ARGUMENT}) + (a_{c,0})_i \cos(\text{ARGUMENT}) \right] \\
& + \Sigma_i \left[(a_{s,1})_i t \sin(\text{ARGUMENT}) + (a_{c,1})_i t \cos(\text{ARGUMENT}) \right] \\
& + \Sigma_i \left[(a_{s,2})_i t^2 \sin(\text{ARGUMENT}) + (a_{c,2})_i t^2 \cos(\text{ARGUMENT}) \right] \\
& + \ldots \\
Y = {} & -0.006951 - 0.025896t - 22.4072747t^2 \\
& + 0.00190059t^3 + 0.001112526t^4 + 0.0000001358t^5 \\
& + \Sigma_i \left[(b_{c,0})_i \cos(\text{ARGUMENT}) + (b_{s,0})_i \sin(\text{ARGUMENT}) \right] \\
& + \Sigma_i \left[(b_{c,1})_i t \cos(\text{ARGUMENT}) + (b_{s,1})_i t \sin(\text{ARGUMENT}) \right] \\
& + \Sigma_i \left[(b_{c,2})_i t^2 \cos(\text{ARGUMENT}) + (b_{s,2})_i t^2 \sin(\text{ARGUMENT}) \right] \\
& + \ldots
\end{aligned}
\tag{23.6}
$$

where $t =$ (TT $-$ 2000 January 1 d 12 h TT) in days/36 525 and ARG stands for various combinations of the fundamental arguments of the nutation theory, including both lunisolar and planetary terms.

Fluctuations of the LOD

The angular velocity, ω_3, fluctuates from its average nominal value, ω (= 7 292 115 $\times$ 10^{-5} rad s^{-1}). Fluctuations, $\Delta\omega_3$, result in variations, ΔLOD in the length of day (LOD) from its conventional value, D = 86 400 SI seconds (i.e., of TAI) and by corresponding changes, ΔT = (UT1 − TAI) of UT1 w.r.t. the International Atomic Time, TAI. These quantities are linked by the following relationships:

$$\Delta\mathrm{LOD}/D = -\Delta\omega_3/\Omega = -\mathrm{d}\Delta T/\mathrm{d}t, \tag{23.7}$$

$$\Delta T = [\mathrm{UT1} - \mathrm{TAI}](t) = [\mathrm{UT1} - \mathrm{TAI}](t_0) + (\Omega)^{-1} \int_{t_0}^{t} \Delta\omega_3 \ \mathrm{d}t. \tag{23.8}$$

Changes in the angular speed of rotation of the Earth reach 10^{-7} in relative value; they are known since the mid-twentieth century. The main components are (1) a secular variation, attributed to tidal dissipation and conservation of global momentum of the Earth-Moon system, which creates an increase of about 2 ms/century in the length of days; (2) decadal variations that are mainly attributed to core-mantle coupling effects; (3) seasonal variations due to exchange of angular momentum between the solid Earth and surface fluids; and (4) variations due to zonal tides. There are also diurnal and sub-diurnal variations of lower amplitude, mainly due to the oceanic tides. Only a small part of these variations can be represented by a model that allows for prediction over a period longer than 1 month. The majority of changes are deduced from observations. Note that the secular variation can only be obtained from observations over a very long period so that, at present, only former optical astrometry observations can contribute to this determination.

The Observations of Earth's Rotation

The observations used for monitoring the Earth's rotation are coordinated at the international level by the IERS with the IVS, ILRS, IGS, and IDS acting as IERS Technical Centers (TC) in charge of providing individual EOP series and combined solutions for each observational technique. The reference series of the IERS, denoted C04, provides daily values since 1962 of x, y, X, Y, UT1–UTC and changes in the duration of the day, LOD, with their uncertainties. These values result from a combination of operational EOP series, each of them associated with one of the IERS techniques (Gambis 2004; Bizouard and Gambis 2008). The IERS Rapid Service/Prediction Center also uses the meteorological predictions of variations in Atmospheric Angular Momentum to aid in the prediction of near-term UT1–UTC changes (Stamatakos et al. 2009). Accuracy for IERS estimates at a daily resolution is currently of the order of a few microseconds (μs) in UT1 and of a few tens of

microarcseconds (μas) for polar motion and precession-nutation quantities. Such an accuracy is a thousand times better than that obtained 35 years ago by optical astrometry, for a time resolution, which may also be, in some cases, a thousand times finer.

Principles of the VLBI and GNSS Techniques for Measuring Earth Rotation

The VLBI Technique and the International VLBI Service for Geodesy and Astrometry (IVS)

Very Long Baseline Interferometry (VLBI) is a technique that allows the observation of extragalactic radio sources from a baseline consisting of two Earth-based antennas separated by several hundreds to several thousand kilometers; each antenna is equipped with high sensitivity receivers, a local oscillator with very high stability during the period of observation (hydrogen maser clock) and a recorder. A measurement consists of the simultaneous recording of the arrival time of a radio wave front emitted by a radio source, at the two antennas (see Fig. 23.10). The processing of the signals is then performed by a correlator, which provides the *time delay*, i.e., the difference between the two arrival times, and the time derivative of that delay. The observations are conducted at two frequencies bands, S band (2.3 GHz) and X band (8.4 GHz) in order to reduce the effect due to the ionosphere.

Originally developed as a radio astronomical technique for high-resolution mapping of the structure of distant quasars, VLBI was implemented in the 1970s as part of geophysical programs, such as the NASA Crustal Dynamics Project designed to monitor tectonic plate motion and Earth's rotation, with the quasars as the reference sources. These measurements have provided, as by-products, the coordinates of the observed sources, as well as geodetic coordinates of the antennas,

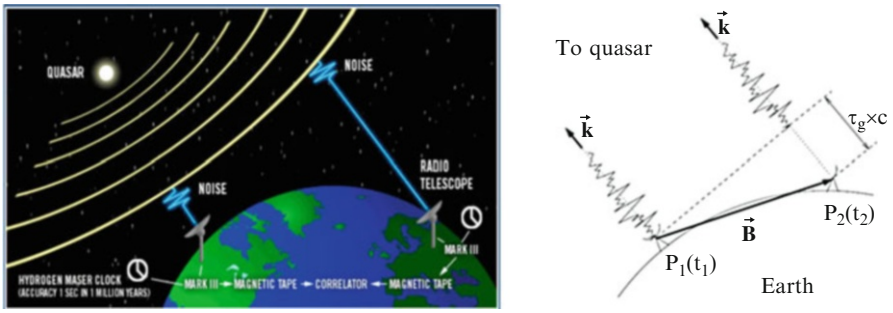


Fig. 23.10 Schematic representation of a VLBI observation (*left*) and the geometric delay (*right*) (source: *left*, IVS home page; *right*, Yaya 2002)

which have been shown to be of very high quality, with a level of uncertainty of one centimeter for the coordinates of the antennas and better than one mas for the sources. Specific programs aimed at obtaining high-resolution measurements of the positions of extragalactic radio sources were then undertaken in the USA, in cooperation with the international scientific community. These programs involved repeated observations of the same sources over long periods of time that led to the first realization of the ICRF (Ma et al. 1998). The International VLBI Service for Geodesy and Astrometry (IVS) was created in 1999 for the coordination of the global VLBI components and resources on an international basis. It is a service of the International Association of Geodesy (IAG), also reporting to the IAU. Its mission is to provide products for the realization of the Celestial Reference Frame (CRF) through the positions of quasars; to deliver products for the maintenance of the terrestrial reference frame (TRF), such as station positions and their changes with time; and to generate products describing the rotation and orientation of the Earth; it is now part of the Global Observing Geodetic System (GGOS) of the IAG (Schuh and Behrend 2012).

Basic Principle of VLBI Observations

A VLBI measurement, or VBI delay, is the time difference between the arrival times, t_1 and t_2, of the radio wave front emitted by a distant quasar, Q, at date t_0 and received at the two Earth-based antennas, P_1 and P_2. The *geometrical delay*, τ_G, is a measure of the optical path difference between the antennas (see Fig. 23.10). As the two antennas point simultaneously to the direction of Q, τ_G is the scalar product of the baseline vector $\mathbf{b}$ and the unit vector $\mathbf{k}$ in the direction of Q, divided by the speed of light c:

$$\tau_G = -\mathbf{k} \cdot \mathbf{b}/c. \tag{23.9}$$

This scalar product is dependent on the length of the baseline and its orientation relative to the Earth, as well as on the direction of the source relative to the Earth; it therefore depends on the Earth's orientation.

In addition to the geometrical delay τ_G plus the relativistic delay correction $\Delta\tau_{rel}$, an observed VLBI delay, τ_{obs}, includes the contributions of various effects that are not a priori known with the necessary accuracy to allow a direct determination of the geometric delay. These contributions come from (1) the mis-synchronization of the station reference clocks that is variable with time, $\Delta\tau_{clock}$; (2) physical effects, such as the propagation delays through the troposphere, $\Delta\tau_{trop}$, and the ionosphere, $\Delta\tau_{iono}$; and (3) the sum, ε, of the un-modeled delays plus the delay measurement noise. A simplified expression of τ_{obs} is as follows:

$$\tau_{\text{obs}} = \tau_{\text{G}} + \Delta\tau_{\text{rel}} + \Delta\tau_{\text{clock}} + \Delta\tau_{\text{trop}} + \Delta\tau_{\text{iono}} + \varepsilon. \tag{23.10}$$

Observed VLBI delays are currently precise to a few picoseconds (10^{-12} s).

Basic Principles of the EOP Estimation from VLBI Observations

The theoretical expression, τ_{th}, of the VLBI delay can be developed as a function of the parameters to be estimated. Such a development requires applying BCRS-to-GCRS transformations to the coordinates of the vectors **k** (i.e., the direction of the quasars), ITRS-to-GCRS transformations (based on the EOP, cf. Eq. (23.4)) to the coordinates of the vectors **b** (i.e., the direction of the antennas), and time transformations providing VLBI delays in TCG. The computation of τ_{th} at each date of the observations is based on a priori values for the parameters as well as on models for clock errors, effects of tectonic movements, Earth and oceanic tides, propagation delays, relativistic effects, etc. (cf. IERS Conventions 2010), each of these models being functions of parameters to be estimated by the observations.

The determination of the corrections to the a priori values of the parameters (ITRS station coordinates, EOP, and BCRS quasar coordinates) from a set of observed VLBI delays requires, as it is the case for any estimation of this kind of parameters, a global statistical adjustment based on the comparison between the observed and computed VLBI delays, which is generally based on the least squares method or Kalman filter. Reducing the formal uncertainties in the estimated parameters requires the use of a large number of delay measurements. The direct access to the IRCF is made via the directions of distant extragalactic sources; this allows VLBI to be a unique technique for the accurate determination of quantities that refer to the ICRF, such as UT1 and the celestial pole offsets. The access to the ITRF is made via the coordinates of the VLBI baselines, which depends on the quality of the global VLBI network. The accuracy of the estimated parameters depends on the distribution on the sky of the observed set of quasars as well as on the geographic distribution of the global network of antennas.

Several software packages have been developed for VLBI data analysis, the standard least squares fit approach being summarized in Fig. 23.11.

IVS Observation Programs and IVS EOP Series

The current IVS observing programs include (cf. Schuh and Behrend 2012) three-and-a-half 24-h sessions per week from a network of 8–10 stations and seven daily 1-h intensive sessions per week from a network of 2–3 stations. The daily "Intensives" are devoted to the determination of UT1–UTC, whereas the unequally

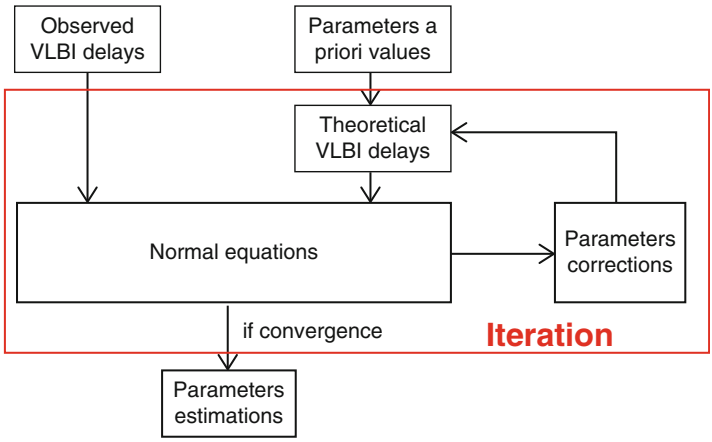


Fig. 23.11 Schematic approach of a VLBI analysis software used for EOP estimation (source: Yao 2013)

spaced 24-h sessions yield the full set of parameters (i.e., for the EOP: x, y, UT1–UTC, x-rate, y-rate, LOD, dX, dY with respect to IAU 2006/2000A precession-nutation model).

The IVS Combination Center provides (cf. IVS website) combined EOP results on the basis of SINEX (Solution Independent Exchange) normal equations produced by the analysis centers, for two kinds of solutions: (1) the quarterly combined solution based on all available sessions from 1984, which is updated every three months, and (2) the rapid combination solution based on all available intensive UT1 sessions from 2005, which is updated twice a week. In addition, specific IVS programs and analyses aim at providing products for the realization of the ICRF and maintenance of the ITRF. A few continuous VLBI campaigns (named "Cont campaigns") were also organized since 1994 to support high-resolution studies.

Nothnagel et al. (2015) have assessed the strengths (S) and weaknesses (W) of the geodetic and astrometric VLBI as follows:

1. (S): It is the only technique for the determination of UT1–UTC and celestial pole offsets and for GNSS polar motion validation; it is able to provide simultaneously the five EOP and UT1–UTC in near real time; and there is a long stable history for all the EOP and a capability for high-frequency resolution.
2. (W): The results depend on the network and there is no redundancy; there is an unequal, noncontinuous sampling; the latency is still too high; there is a dependency on station stability (position/equipment); the same correlation/fringing program is used; there is only a few data analysis programs; and the IVS UT1 Intensives are not combined rigorously.

A new system, called VGOS (VLBI2010 Global Observing System), of which development work has been under way since 2005, is expected to replace the current S/X system in the next several years. This will allow greater observation density, more precise delay observables, better station network, lower latency,

continuous sampling, increased robustness, less vulnerable to station-specific events, etc., which will provide a more accurate determination of the Earth orientation.

The GNSS Technique and the GNSS International Service (IGS)

The Global Navigation Satellite System (GNSS) is a satellite-based global radio-navigation system referring to a constellation of satellites and their ground stations. A constellation is composed of groups of eight to ten satellites on a few circular orbits having the same inclination, at an altitude of approximately 20,000 km (corresponding to a satellite orbital period of approximately 12 h). The satellites are equipped with very stable atomic clocks that are synchronized to each other and to ground clocks and are continuously transmitting coded signals at precise intervals. Two carrier signals are necessary for self-calibration of the ionospheric delay. Coordinating the signal data from four or more satellites enables the receiver to determine its position and its clock offset with respect to the GNSS time.

The first GNSS system was the Global Positioning System (GPS), originally called Navigation System by Timing and Ranging (NAVSTAR), was launched in the USA in 1973, and became fully operational in 1995. It now consists of a constellation of about 30 operational satellites that orbit the Earth in six orbital planes spaced 60° apart and inclined at 55° relative to the equator at an altitude of approximately 20,200 km (see Fig. 23.12). They generate a radio signal on two well-defined frequencies L1 (1.57542 GHz) and L2 (1.2276 GHz), which are modulated in phase by pseudorandom codes (military or civilian). Several improvements to the GPS service have been implemented since its deployment, including a

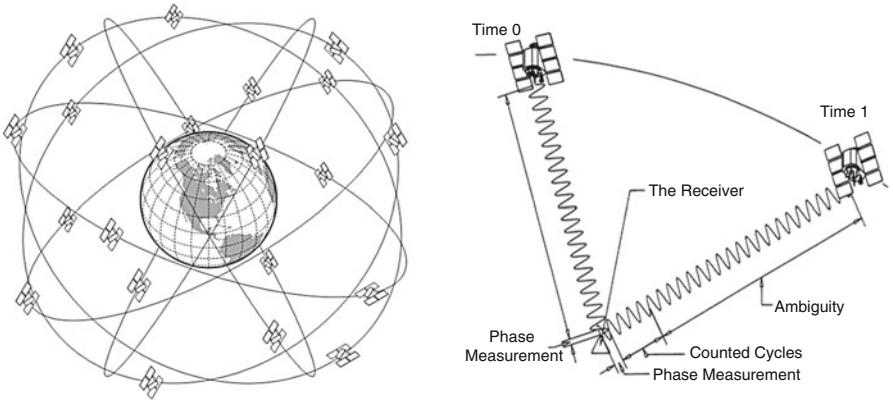


Fig. 23.12 GPS constellation (*left*) and diagram of the phase measurement (*right*) (source: *left*: https://www.e-education.psu.edu/geog862/node/1770; *right*: Porqueras 2012)

new generation of GPS satellites, new signals for civil use, and increased accuracy and integrity for all users.

The International GNSS Service (IGS) was established in 1994 for the coordination of global observations and analysis of GPS and then GNSS observations,[3] leading to many scientific applications on reference systems, positioning, navigation, clock synchronization, etc. IGS is a service of the IAG and is part of GGOS and operates in close cooperation with the IERS. Its mission is (cf. IVS website) to provide, on an openly available basis, the highest-quality GNSS data, products, and services in support of the terrestrial reference frame; Earth observation and research; Positioning, Navigation, and Timing (PNT); and other applications that benefit the scientific community and society. The IGS tracking network consists of about 400 permanent stations with a worldwide geographic distribution.

Basic Principle of GNSS Observations

A basic GNSS measurement is the time of signal propagation between the satellite and the station, calculated by correlating the satellite signal and the receiver signal. This time can be converted into a *pseudo-range* measurement, ρ, by multiplying it by the speed of light. The GPS carrier-phase mode, which requires the use of geodetic GPS receivers, uses both the L1 and L2 carrier frequencies instead of the codes transmitted by the satellites to obtain *carrier-phase* measurements, Φ. This is able to provide very high-precision range, but with an ambiguity in the number of whole carrier cycles (see Fig. 23.12).

GNSS measurements, similar to those of VLBI, must be corrected for contributions from clock errors, physical effects, and noise. Simplified equations can be expressed as:

$$\begin{aligned} \rho_{\mathrm{obs}} &= \rho_{\mathrm{g}} + \rho_{\mathrm{rel}} + \rho_{\mathrm{clocks}} + \rho_{\mathrm{trop}} + \rho_{\mathrm{ion}} + \varepsilon \\ \lambda\Phi_{\mathrm{obs}} &= \rho_{\mathrm{g}} + \rho_{\mathrm{rel}} + \rho_{\mathrm{clocks}} + \rho_{\mathrm{trop}} + \rho_{\mathrm{ion}} + \varepsilon + \lambda N \end{aligned} \tag{23.11}$$

where ρ_{g} is the geometrical distance between the transmitter at the emission date t_1 and the receiver at the reception date t_2, and ρ_{rel} is the relativistic correction. The terms, ρ_{clocks}, ρ_{trop}, ρ_{iono}, and ε are due to clock errors, tropospheric and ionospheric contributions, and un-modeled effects plus the noise; λ is the carrier wavelength and λN the ambiguity parameter. Observed GPS phase measurements currently provide precision to a few millimeters.

[3]The previously adopted name, "International GPS Service," was changed to "International GNSS Service" on 2005, due to further expansion of IGS, integrating data from the Russian GLONASS system and from future GNSS, such as the Chinese Beidou or European Galileo.

Basic Principles of the EOP Estimation from GNSS Observations

The theoretical expression of GNSS measurements can be developed as a function of the parameters to be estimated by a global statistical adjustment; this must be preceded by the determination of station clock biases based on the pseudo-range observations in order to compute the satellite positions. The computation of the theoretical quantity at each date of the observations requires in particular applying a GCRS-to-ITRS coordinate transformation (based on the EOP, cf. Eq. (23.4)) to the satellite coordinates as well as time transformations. Similarly to the VLBI case, the computation is based on a priori values for the parameters as well as models corresponding to various effects. A crucial part of GNSS analysis is the high-precision determination of the orbits of the satellites, which is affected by systematic errors in the gravitational and nongravitational forces. As the access to the CRF is provided through these orbits, the GNSS technique cannot provide accurate determination of quantities that refer to the CRF, such as UT1 and the celestial pole offsets (see the next paragraph).

Many software packages have been developed for GNSS data analysis, the standard least squares fit approach being summarized in Fig. 23.13.

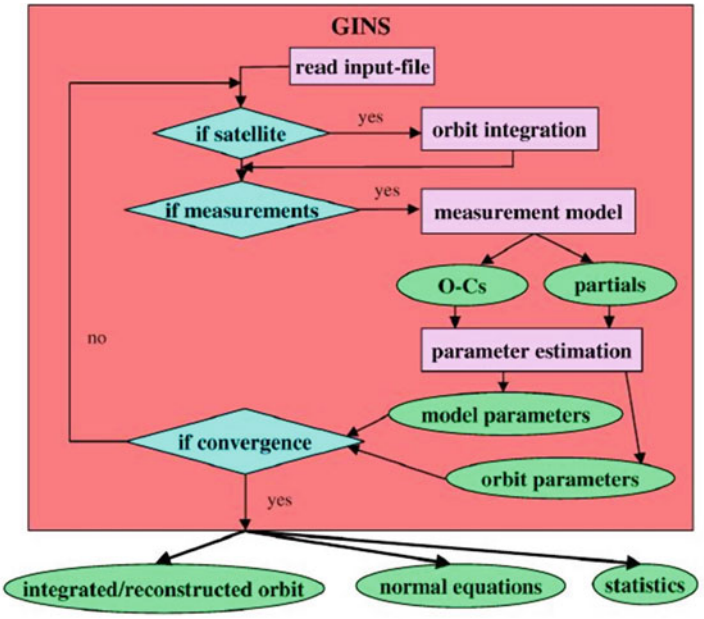


Fig. 23.13 Schematic approach of a GNSS analysis software used for EOP estimation (source: Meyer et al. 2000)

Consequences of the Systematic Errors in the Orbital Elements on the GNSS Determination of the EOP

The principle of determining the EOP from Earth satellite observations comes from the fact that inaccuracies in the EOP will give rise to discrepancies between the computed and observed satellite positions in the GCRS. Such discrepancies can then be used to derive corrections for the EOP. However, the crucial issue is to separate the discrepancies due to corrections to the EOP from those due to deficiencies in the models for the gravitational and nongravitational forces acting on the satellite. The dependence of the orbital elements of the satellite (defining the orientation of the orbital plane of the satellite in the GCRS) on the EOP can easily be obtained by expressing the transformation of the satellite coordinates from the GCRS to the ITRS (cf. Eq. (23.4), Capitaine and Wallace 2007).

The differences between the orbital elements that refer to the ITRF equator and the CIP equator can be expressed as follows (at the first order in x, y, and neglecting the very small s' quantity):

$$\begin{aligned} \Delta i &= x \sin(\mathrm{ERA} - \Omega) - y \cos(\mathrm{ERA} - \Omega) \\ \Delta\Omega \tan i &= -x \cos(\mathrm{ERA} - \Omega) - y \sin(\mathrm{ERA} - \Omega) \\ \Delta u_0 \sin i &= x \cos(\mathrm{ERA} - \Omega) + y \sin(\mathrm{ERA} - \Omega) \end{aligned} \tag{23.12}$$

where Ω and i are the right ascension of the ascending node and inclination of the orbital plane and u_0 is the argument of the latitude of the satellite at the osculating epoch.

Similar relationships can be written for the differences in the orbital elements as functions of the corrections to the precession-nutation quantities and the a priori value of the ERA. The changes in the observed orbital elements of an artificial satellite corresponding to celestial pole offsets (i.e., corrections to the GCRS CIP coordinates, X, Y) and correction to the ERA at date t can be expressed as follows (at the first order in dX, dY, and d(ERA), neglecting the small s quantity, and with Keplerian approximation):

$$\begin{aligned} \Delta i &= -\mathrm{d}X \sin\Omega + \mathrm{d}Y \cos\Omega \\ \Delta\Omega \tan i &= (-\mathrm{d}X \cos\Omega - \mathrm{d}Y \sin\Omega) - \tan i \ \ \mathrm{dERA}, \\ \Delta u_0 \sin i &= \mathrm{d}X \sin\Omega + \mathrm{d}Y \cos\Omega \end{aligned} \tag{23.13}$$

Comparing relations (23.12) and (23.13) shows that, although the effects of the terrestrial and celestial motions of the CIP on the orbital elements of the satellites have similar forms, they appear at different periods. The effect of polar motion gives rise to variations of around a 1-day period (i.e., period of the ERA) in the orbital elements, while the effect of the corrections dX, dY, and d(ERA) appears as slow periodic variations (i.e., period of the ascending node of the orbital plane).

Variations with a period of around 1-day period in the orbital elements can easily be separated from the systematic errors in those elements, while long periodic terms cannot. This, together with the excellent distribution of observations over the period during which the orbit is computed, explains why the GNSS technique is very powerful for the determination of polar motion, while absolute determination of CIP offsets, or ERA, is not possible. There are strong correlations between those latter parameters and the systematic errors in the orbital elements due to deficiencies in the force models. As these correlations depend on the value for the inclination and the longitude of the node of the orbit of the satellite, improvements can be obtained with the combined use of different GNSS (see, e.g., Weber and Englich 2006).

Contribution of GPS to the Determination of UT1 and Celestial Pole Offsets

We limit this paragraph to GPS, which is currently the only GNSS technique providing high-precision EOP on a regular basis. For the reasons explained above, GPS has a very good potential to determine polar motion and its time variations (pole rates), but it cannot determine the absolute UT1 or celestial pole offsets. Different approaches that are briefly summarized in the next paragraph have been followed so that GPS can contribute to the UT1 estimate despite the systematic deficiencies.

The possibility of determining a UT1-like quantity, called UTGPS, from analysis of GPS orbital planes was studied by Kammeyer (2000). Another approach is based on the fact (see Senior et al. 2010; Gambis and Luzum 2011) that, on time scales up to about 20 days, the orbital systematic errors remain limited. Therefore, GPS can offer the potential of estimating the very short periodic variations of UT1, which can be used for densification of UT1 derived by VLBI. The IERS Earth Orientation Center currently makes use of the integrated LOD(GPS) series, corrected for the systematic drift and calibrated by the IVS solution, to give a so-called UT(GPS) solution, which has been shown to be significantly less noisy than the IVS intensive combined solution (Gambis 2014). Therefore, UT(GPS) is entering the IERS C04 solution with a significant weight, contributing to a fair regularization in the stability of the C04 (Bizouard and Gambis 2008; Gambis and Luzum 2011). The LOD determinations made by analysis of GPS observations can also be a useful part of rapid combinations (Gambis and Luzum 2011).

Specific studies have also been made on the possible contribution of GPS to the determination of nutation. Similarly to the LOD approach, the rates of the celestial pole offsets are estimated, assuming that the perturbing forces on the satellite can be sufficiently well modeled during the 3-day period of the orbit computations (Rothacher 1999; Rothacher et al. 1999) and using a priori VLBI values. Based on this method, Weber and Rothacher (2001) provided a revised set of nutation

amplitudes from 6 years of GPS data. Corrections to the IAU 2000 nutation have also been obtained by the combination of VLBI-based celestial pole offsets with the GPS-based celestial pole offset rates (Vondrák et al. 2003). Then, encouraging results have been obtained in a work investigating the potential of GNSS observations, independently of VLBI, for CIP offset estimations (Yao 2013). This showed that the GNSS technique alone (i.e., without any use of VLBI estimates) has the potential of estimating simultaneously polar motion and the time derivatives of the CIP offsets (see Fig. 23.17) and of the ERA. In order to do this, the satellite orbit integration must be performed in an inertial reference system minimizing the influence of the a priori values for precession-nutation X, Y, and ERA that GPS cannot determine directly. Corrections to the amplitudes of the short periodic terms of the IAU 2000 nutation have then been derived for a 3-year series of these quantities.

IGS Observing Programs and IGS EOP Series

The GPS stations provide observations in the Receiver Independent Exchange (RINEX) format, some of them every 30 s, others at each second. These observations are collected, combined, and delivered by IGS, which provides GNSS orbits, tracking data, and other high-quality GNSS data, as well as near real-time, online data products and their characteristics (accuracy, latency, etc.). Four kinds of IVS products are provided: ultra-rapid (predicted half), ultra-rapid (observed half), rapid, and final.

The mean RMS of the final IGS products are currently of the order of 2.5 cm in the three geocentric coordinates of the satellites, 75 ps in the satellite and station clocks, a few mm in the geocentric station coordinates, and about 1 mm/year for the station velocities. The characteristics of the four kinds of IVS EOP products are shown in Fig. 23.14.

Comparisons and Results

Comparisons

Table 23.1, based on tables from Gambis and Luzum (2011) and the website of the IERS EOP Product Center (http://hpiers.obspm.fr/eop-pc/products/combined/verif.html), provides a comparison of the characteristics of the official IVS and IGS series of EOP; the accuracy of the solution (column 4) has been estimated by comparison to the combined solution of the IERS Rapid/Prediction Service at USNO or the IERS C04 solution of the IERS Product Center at Paris Observatory.

Polar Motion (PM) Polar Motion Rates (PM rate) Length-of-day (LOD)

Type		Accuracy	Latency	Updates	Sample Interval
Ultra-Rapid (predicted half)	PM	~200 μas	real time	at 03, 09, 15, 21 UTC	daily integrations at 00, 06 12, 18 UTC
	PM rate	~300 μas/day			
	LOD	~50 μs			
Ultra-Rapid (observed half)	PM	~50 μas	3 - 9 hours	at 03, 09, 15, 21 UTC	daily integrations at 00, 06, 12, 18 UTC
	PM rate	~250 μas/day			
	LOD	~10 μs			
Rapid	PM	~40 μas	17 - 41 hours	at 17 UTC daily	daily integrations at 12 UTC
	PM rate	~200 μas/day			
	LOD	~10 μs			
Final	PM	~30 μas	11 - 17 days	every Wednesday	daily integrations at 12 UTC
	PM rate	~150 μas/day			
	LOD	~10 μs			

Note 1: 100 μas = 3.1 mm of equatorial rotation; 10 μs = 4.6 mm of equatorial rotation.

Note 2: The IGS uses VLBI results from IERS Bulletin A to partially calibrate for LOD biases over 21-day sliding window, but residual time-correlated LOD errors remain.

Fig. 23.14 Characteristics of the four kinds of IGS EOP products (source: IGS home page)

Table 23.1 Characteristics (time resolution and estimated accuracy) of the combined IVS and IGS series of EOP (source: Gambis and Luzum 2011; Gambis et al. 2014)

Technique	Since	Parameter	Time resolution	Accuracy
VLBI standard	1981	UT1	3–4 days	5 μs
VLBI intensive	1981	UT1	1 day	15 μs
VLBI	1981	LOD	3–4 days	15 μs d^{-1}
GPS	1993	LOD	1 day	10 μs d^{-1}
VLBI standard	1981	Polar motion	3–4 days	80 μas
GPS	1993	Polar motion	1 day	30 μas
VLBI standard	1981	CIP offsets	3–4 days	40 μas

A Few Illustrations of EOP Results Based on VLBI and GPS Observations

The scientific importance of the VLBI- and GPS-based EOP determination is illustrated by three figures. The first one (Fig. 23.15) shows the strong correlation between the axial component of Earth's rotation, as determined by the IERS combination of VLBI and GPS results, and the atmospheric and oceanic contributions to the Earth's angular momentum. This is fundamental for a better understanding of the global Earth's dynamics.

The second one (Fig. 23.16) shows a comparison over 30 years between VLBI determinations of the CIP offsets and the current IAU precession-nutation model, which clearly reveals the existence of the free core nutation (FCN) of a 430-day period in space, as well as its amplitude variations with time. This is fundamental for a better understanding of the Earth's mantle-core dynamics.

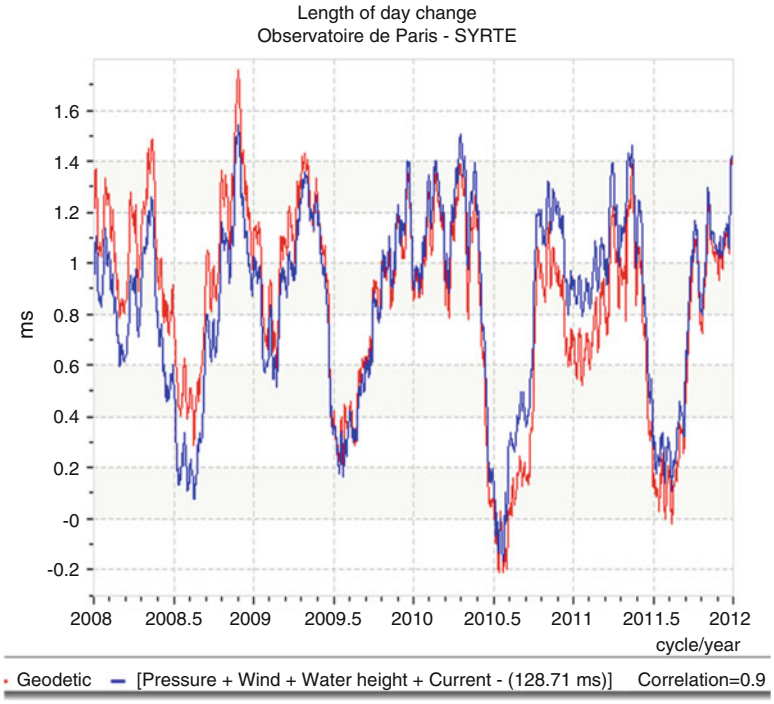


Fig. 23.15 Comparison, from 2008 to 2012, between the geodetic and geophysical axial components of the excitation function (The term "excitation function" is that used in the "Euler-Liouville" equations designed for computing variations of Earth rotation from angular momentum changes and torques acting on the Earth.), the geodetic component being derived from the IERS C04 ΔLOD solution resulting from a VLBI/GPS combination and the geophysical component being derived from the atmospheric and oceanic contributions to the Earth's angular momentum (source: http://hpiers.obspm.fr/eop-pc/products/)

The third one (Fig. 23.17) shows the short-term periodic variations of the CIP offset rates over a period of 100 days determined by GPS only. This can potentially improve our understanding of the short-term nutations.

Summary

International resolutions on space-time reference systems and Earth's rotation have been adopted from 1997 to 2009 and have provided the modern bases for the determination of Earth orientation, including improved celestial and terrestrial reference frames, concepts, definitions and models, as well as a new paradigm. All the astro-geodetic techniques have now to be considered in this new framework.

VLBI and GPS techniques are the two modern techniques currently contributing the most to the determination of Earth orientation. The differences in their principle

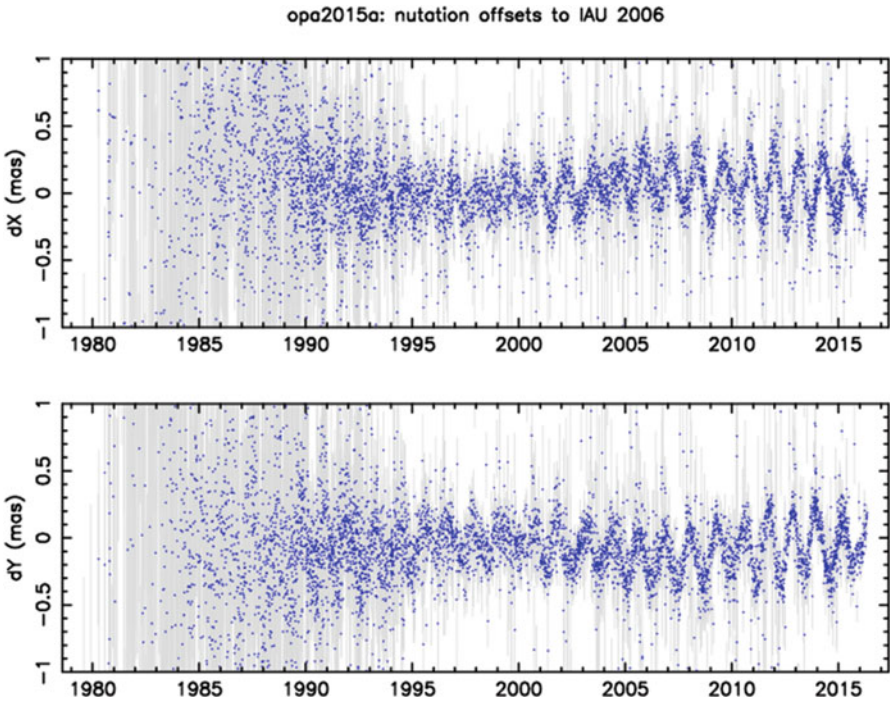


Fig. 23.16 Series of VLBI celestial pole offsets (corrections to the IAU 2006/2000 precession-nutation model) (source: http://ivsopar.obspm.fr)

of observation and their access to the terrestrial and celestial reference systems give these techniques different potentials to determine the EOP. VLBI is able to simultaneously provide the five EOP and is unique in determining UT1 and the celestial pole offsets. The time series of the five VLBI-based EOP are available over a period of more than 30 years. In contrast, GPS has currently a major contribution to the determination of the pole coordinates x, y and also provides LOD estimates on a regular basis. The time series of these quantities are available over a period of about 20 years. GPS cannot provide absolute UT1 or celestial pole offsets, but it is able to estimate the very short periodic variations of UT1, which can be used for densification of UT1 derived by VLBI. It is also able to provide the rates of celestial pole offsets that can be used to estimate short periodic nutation.

The VLBI and GPS modernization programs of which development work is under way are expected to significantly improve the accuracy and time resolution of these two techniques for EOP determination. This, together with the ongoing developments of other GNSS and their interoperability, and future combination of those techniques will offer new opportunities for the determination of Earth rotation variations, from long-term variations to very short periodic terms.

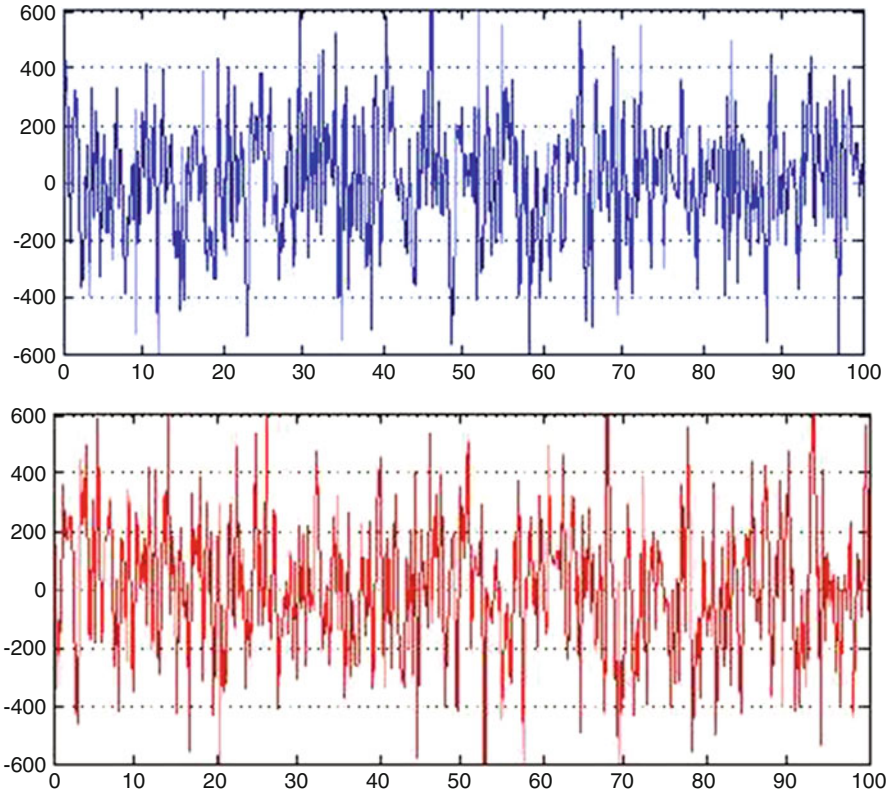


Fig. 23.17 Series of CIP offsets rates, d*X*dot, d*Y*dot, over a 100-day interval from 01/01/2009 (units: μas/d) (source: Yao 2013)

References

Z. Altamimi, P. Rebischung, L. Métivier, X. Collilieux, *Key Performance Indicators of the ITRF2014 Products, Geophysical Research Abstracts*, vol. 18, EGU2016-6995 (2016)

C. Bizouard, D. Gambis, *The Combined Solution C04 for Earth Orientation Parameters Consistent with International Terrestrial Reference Frame*. http://hpiers.obspm.fr/eoppc/eop/eopc04/C04.guide.pdf (2008)

N. Capitaine, B. Guinot, D.D. McCarthy, Definition of the celestial ephemeris origin and of UT1 in the international reference frame. Astron. Astrophys. **355**, 398 (2000)

N. Capitaine, B. Guinot, J. Souchay, A non-rotating origin on the instantaneous equator: definition, properties and use. Celest. Mech. Dyn. Astron. **39**, 283 (1986)

N. Capitaine, P.T. Wallace, J. Chapront, Expressions for IAU 2000 precession quantities. Astron. Astrophys. **412**, 567–586 (2003)

N. Capitaine, P.T. Wallace, High precision methods for locating the celestial intermediate pole and origin. Astron. Astrophys. **450**, 855 (2006)

N. Capitaine, A. Andrei, M. Calabretta, V. Dehant, T. Fukushima et al., Proposed terminology in fundamental astronomy based on IAU 2000 resolutions, in *Highlights of Astronomy*, vol 14 (2007), pp. 474–475

N. Capitaine, P.T. Wallace, The transformation between the terrestrial and celestial reference systems: needs and potential of GPS and Galileo, in *Proceedings of the ESA 1st Colloquium Scientific and Fundamental Aspects of the Galileo Programme*, ESA on CD, 2007

A.L. Fey, D. Gordon, C.S. Jacobs, C. Ma, R.A. Gaume, et al., The second realization of the international celestial reference frame by very long baseline interferometry. Astron. J. **150**, 58 (2015)

D. Gambis, Monitoring Earth orientation using space-geodetic techniques: state-of-the-art and prospective. J. Geod. **78**(4–5), 295–303 (2004)

D. Gambis, B. Luzum, Earth rotation monitoring, UT1 determination and prediction. Metrologia **48**(4), S165 (2011)

D. Gambis, Contribution of GPS to Earth Orientation Monitoring, IGS Workshop Celebrating 20 Years of Service 1994–2014 (2014)

D. Gambis et al., IERS Annual Report 2014, Report from the Earth Orientation Center (2014)

B. Guinot, Basic problems in the kinematics of the rotation of the Earth, in *Time and the Earth's Rotation*, ed. by D.D. McCarthy, J.D. Pilkington (D. Reidel Publishing Company, Dordrecht, 1979), pp. 7–18

J. Hilton, N. Capitaine, J. Chapront, J.M. Ferrandiz, A. Fienga, Report of the International Astronomical Union Division I Working Group on precession and the ecliptic. Celest. Mech. Dyn. Astron. **94**(3), 351–367 (2006)

IERS Conventions, in *IERS Technical Note 36*, ed. by G. Petit, B. Luzum (Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, 2010), 179 pp., ISBN 3-89888-989-6

IAU 1997, in *Transactions of the IAU Vol. XXIII B, Proceedings of the 23rd General Assembly*, Kyoto, Japan, 18–30 August 1997, ed. by J. Andersen (1999)

IAU 2000, in *Transactions of the IAU Vol. XXIV B, Proceedings of the 24th General Assembly*, Manchester, UK, 7–18 August 2000, ed. by H. Rickman, ASP (2001)

IAU 2006, in *Transactions of the IAU, Volume XXVI B*, ed. by K.A. van der Hucht (Cambridge University Press, 2008)

IAU 2009, in *IAU Transactions XXVII B, Proceedings of the 27th IAU GA, General Assembly*, Rio de Janeiro 2009, ed. by I.F. Corbett (Cambridge University Press, 2010)

IUGG 2003, in *Resolutions of the 23rd IUGG General Assembly Sapporo Japan*, June 30–July 11, 2003. http://www.iugg.org/resolutions/sapporo03.pdf

IUGG 2007, in *Resolutions Adopted by the Council at the 24th IUGG General Assembly*, Perugia, Italy, 2–13 July 2007. http://www.iugg.org/resolutions/perugia07.pdf

IUGG 2011, in *Resolutions Adopted by the Council at the 25th IUGG General Assembly*, Melbourne, Australia, June 26–July 7, 2011. http://www.iugg.org/resolutions/ IUGG Resolutions - XXV GA - Melbourne (English).pdf

A. Nothnagel, D. Behrend, A. Bertarini, J. Boehm, P. Charlot et al., in *Report Summary of IVS 2015 Retreat "IVS Strategy for the next decade (IVS2020)"* , 7–8 October 2015, Penticton, BC, Canada, 2015

P. Kammeyer, A UT1-like quantity from analysis of GPS orbit planes. Celest. Mech. Dyn. Astron. **77**, 241–272 (2000)

S. Klioner, N. Capitaine, W.M. Folkner, T.-Y. Huang et al., Units of relativistic time scales and associated quantities, in *Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis, Proceedings of the IAU Symposium*, 2010, vol. 261, pp. 79–84

C. Ma, E.F. Arias, T.M. Eubanks, A. Fey, A.-M. Gontier, et al., The international celestial reference frame as realized by very long baseline interferometry. Astron. Astrophys. **116**, 516 (1998)

P.M. Mathews, T.A. Herring, B.A. Buffett, Modeling of nutation-precession: new nutation series for nonrigid Earth, and insights into the Earth's interior. J. Geophys. Res. **107**(B4), ETG 3-1 (2002). doi:10.1029/2001JB000390

U. Meyer, P. Charlot, R. Biancale, GINS: a new multi-technique software for VLBI analysis, in *International VLBI Service for Geodesy and Astrometry 2000 General Meeting Proceedings*, ed. by F. Takahashi (2000), pp. 324–328

M. Porqueras, GPS Time and Frequency Transfer Techniques, Navipedia, ESA (2012). http://www.navipedia.net
M. Rothacher, The contribution of GPS measurements to Earth rotation studies, in *Journées 1998 "Systèmes de référence spatio-temporels," Observatoire de Paris*, ed. by N. Capitaine (1999), p. 239
M. Rothacher, G. Beutler, T.A. Herring, R. Weber, Estimation of nutation using the global positioning system. J. Geophys. Res. **104**(B3), 4835–4859 (1999)
H. Schuh, D. Behrend, VLBI: a fascinating technique for geodesy and astrometry. J. Geodyn. **61**, 68–80 (2012). doi:10.1016/j.jog.2012.07.007
K. Senior, J. Kouba, J. Ray, Status and prospects for combined GPS LOD and VLBI UT1 measurements. Artif. Satell. **45**, 57–73 (2010)
M. Soffel, S. Klioner, G. Petit, P. Wolf, S. Kopejkin, et al., The IAU 2000 resolutions for astrometry, celestial mechanics, and metrology in the relativistic framework: explanatory supplement. Astron. J. **126**(6), 2687–2706 (2003)
N. Stamatakos, B. Luzum, B. Stetzler, W. Wooden, E. Schultz, Recent improvements in IERS rapid service/prediction center products, in *Proceedings of the Journées 2008 "Systèmes de Référence spatio-temporels,"* ed. by N. Capitaine (Observatoire de Paris, 2009), pp. 160–163
J. Vondrák, C. Ron, R. Weber, Combined VLBI/GPS series of precession-nutation and comparison with IAU2000 model. Astron. Astrophys. **397**, 771 (2003)
J. Vondrák, N. Capitaine, P.T. Wallace, New precession expressions, valid for long time intervals. Astron. Astrophys. **534**, A22 (2011)
P.T. Wallace, in *Highlights of Astronomy*, vol. 11, ed. by J. Andersen (Kluwer Academic Publishers, Dordrecht, 1998), p. 191
R. Weber, M. Rothacher, A revised set of nutation amplitudes calculated from six years of GPS data, in *AGU Meeting 2001*, abstract G51C-0252 (2001)
R. Weber, S. Englich, Potential contribution of Galileo to the TRF and the determination of ERP, in *Journées 2005 Systèmes de Référence Spatio-Temporels*, ed. by A. Brzezinski, N. Capitaine, B. Kolaczek (Space Research Centre PAS, Warsaw, Poland, 2006), p. 294
K. Yao, Nutation determination by VLBI and GPS techniques, Thèse de doctorat (PhD), Univ. Pierre et Marie Curie (29 avril 2013)
Ph. Yaya, Combinaisons des techniques d'astrométrie et de géodésie spatiale, Thèse de doctorat (PhD), Observatoire de Paris (1er juillet 2002)

List of a Few Useful Websites

IERS website: https://www.iers.org
IGS website: http://www.igs.org
IVS website: http://ivscc.gsfc.nasa.gov

Chapter 24
Status of the Gaia Mission

François Mignard

Abstract In December 2013, ESA successfully launched the Gaia satellite to survey our Galaxy with astrometry, photometry and spectroscopy. The main goal is to investigate the formation and evolution of the Milky Way through the kinematics and physics of its stars. Astrometry of point sources (stars, quasars) is the core of the mission and will lead to positions, proper motions and distances to an unprecedented accuracy down to 20.5 mag for more than 1 billion sources.

I will explain how the instruments on board can achieve this ambitious goal with a scanning satellite and how it depends heavily on a demanding time metrology based on the continuous operation of a Rb clock on board and its daily correlation to ground-based timescale.

I will report on the actual performances seen after the first exploitations and focus on the importance of Gaia for the realisation of the Celestial Reference Frame and its future implication to monitor the Earth's rotation.

Keywords Astrometry • Reference frame • Milky Way • Time synchronisation

F. Mignard (✉)
Observatoire de la Côte d'Azur, Bld de l'Observatoire, Nice 06304, France
e-mail: francois.mignard@oca.eu

E.F. Arias et al. (eds.), *The Science of Time 2016,* Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_24

Chapter 25
Time Synchronization and the Origins of GPS

Richard D. Easton

Abstract The Global Positioning System has been called the first global utility. Yet its origins are relatively obscure; the timing aspect of this PNT system (positioning, navigation, and timing) is less obvious to the general public than are the positioning and navigation aspects. The slogan of the first head of the GPS Joint Program Office, Col. Brad Parkinson, was, "Drop 5 bombs in the same hole. . .and don't you forget it!" (Parkinson and Powers, Part 2: The Origins of GPS, Fighting to Survive. *GPS World*, June 2010). This article will demonstrate that, contrary to a common assertion in the literature, time synchronization was anticipated by some of the GPS pioneers.

Keywords GPS • Time synchronization

Description

GPS is run by the US Air Force. It has three segments. The **space segment** currently uses 31 satellites; limitations in the ground segment software restrict it to actively using at most 32 satellites, in six evenly spaced constellations (60° apart). The satellites are in close to 12 h circular orbits inclined at 55°. The last of the GPS 2 satellites have been launched, and the new GPS 3 satellites are scheduled to start launching in 2018. GPS 3 currently requires the new ground control system, OCX, to be up and running. However, OCX is behind schedule and over budget. One possibility, which the Air Force is studying, is modifying the current ground system to allow it to control these new satellites. The **ground segment** updates the clocks and orbits of the satellites. The **user segment** is the GPS receiver. The receiver's three-dimensional position and time are obtained when signals from four GPS satellites are received. GPS is a passive system; the receiver does not need to make a transmission. The signal sent from the satellite includes the time of transmission; if the receiver has a synchronized clock, the time difference reveals

R.D. Easton (✉)
Independent Scholar, International Bureau for Weights and Measures, Dallas, TX, France
e-mail: winnetkaelm@comcast.net

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_25

the distance between the satellite and the receiver. With four satellites in sight, the receiver's three-dimensional position and clock synchronization can be computed.

GNSS

GPS was the first global navigation satellite system to be developed. The first GPS Block 1 test satellite was launched in 1978. Three other systems either are or intend to be global in nature. The Soviet Union launched its first GLONASS satellite in 1982. It fell on hard times after the breakup of the Soviet Union but has been built back to 25 satellites. The European Galileo system and the Chinese BeiDou system are being rapidly developed. Regional systems are being built by India (NAVIC (Navigation with Indian Constellation) or IRNSS (Indian Regional Navigation Satellite System)) and Japan (QZSS (Quasi-Zenith Satellite System)). These use satellites in geosynchronous orbits at the longitudes of these countries. Used in combination with GPS or other systems, they improve navigation in these countries. They are particularly helpful in large cities such as Tokyo where the tall buildings can block the signals from GPS satellites.

Was Time Synchronization Anticipated in the Creation of GPS?

There are numerous sources which assert that time synchronization was not anticipated in GPS. Thus, Don Jewell, a longtime Air Force officer who worked on GPS, wrote:

> You need to know that in those days [the late 1970s], GPS was seen as a military system only, and there were no thoughts of granting public, much less worldwide, free access to the GPS signals. It was basically a military en route navigation system, and not much was thought about it beyond that. Certainly there was no serious consideration given to GPS becoming the de facto world standard for time. (Jewell 2008)

Similarly, a 2009 Ph.D. thesis on Galileo stated, "Somewhat unexpectedly, the precise timing information transmitted by GPS satellites was quickly incorporated into many inventive applications that were not related to navigation" (Gleason 2009). Nunzio Gambale [CEO and co-founder of Locata Corporation] stated, "There was never an intention for it [GPS] to be used for infrastructure, for timing" (Milner 2016).

Exhibit 1 shows eight US pre-GPS space-based navigation proposals. They are discussed in detail in my book (Easton and Frazier 2013). Three of these proposals were done by individuals or groups who previously tracked Sputnik. They were George Weiffenbach and William Guier from the Applied Physics Lab, Roy Anderson from General Electric, and Roger Easton from the Naval Research

Laboratory. They then inverted the problem. Instead of using one or more ground stations to track a satellite, they used one or more satellites to provide the position of a receiver. The 1973 synthesis of the last two systems, Timation and 621B, gave us GPS.

Weiffenbach and Guier tracked Sputnik 1 in October 1957 by measuring the Doppler shift in its signal. The following March, they began work on what became Transit, the first space-based navigation system. Roy Anderson, at the Schenectady, NY, office of General Electric, also tracked Sputnik. Later, he tracked Pioneer 4 and had an amusing story about how the media distorts stories:

> Three tracking stations demonstrated the ability to track Pioneer 4 to the great distance: Jodrell Bank in England with its 150-foot diameter antenna, the Jet Propulsion Laboratory (JPL) at Goldstone Lake, California with its 85-foot antenna, and a temporary setup at the GE Research Laboratory, Schenectady, with an 18-foot diameter parabolic antenna. ... There was immense media interest in our effort. Pioneer 4 was the first object to escape Earth's gravity. We were besieged with phone calls at all hours of the day and night. With our small antenna we were seen as David against Goliath. On the morning of 6 March, the signal was weak and intermittent. Finally, search as we could, we could no longer get a lock on it. In mid-morning, a reporter called and said, "JPL announced that they lost a signal. Do you still have it?" "No." "When did you lose it?" "I don't know exactly?" "Can we say that at 10:25 you said that you lost the signal?" It was 10:25. "Yeah." By 10:27 the whole world was informed that GE had tracked the space probe farther than JPL. I went to a newspaper office and asked them to publish a disclaimer. They were not interested. JPL was not pleased (Anderson 2008).

Roger Easton, of the Naval Research Laboratory, was a coauthor of the 1955 proposal which was selected to launch the first American satellite (the project was named Vanguard). It mentioned the use of satellites for navigation, "It would also be possible to determine the absolute longitudes and latitudes by observation of the satellite. Such observations would also yield the height of the observer above the center of the earth" (A Scientific Satellite Program 1955). He helped design the Minitrack system which was intended to track the Vanguard satellite. When Sputnik was launched, modifications were needed to Minitrack since Sputnik transmitted at 20 MHz and 40 MHz, whereas the International Geophysical Year standards specified 108 MHz (McLaughlin and Lomask 2009). In January 1958, before the first American satellite was launched, Explorer 1, he conceived what became the Naval Space Surveillance System, later the Air Force Space Surveillance System, to track Soviet spy satellites which would not emit a signal most of the time. He started the Timation (time navigation) program in 1964. This partially arose out of the need to synchronize the clocks in the two transmitter stations in Texas. We see a symbiotic relationship where space tracking improves the accuracy of satellite navigation systems such as GPS which in turn improves the accuracy of the space tracking systems. The last two Timation satellites, renamed NTS-1 and NTS-2 (Navigation Technology Satellite), carried the first rubidium and cesium atomic clocks into orbit.

Time Synchronization Was Foreseen

Roy Anderson wrote in 1964 that his system would provide worldwide synchronization to approximately 1 microsecond (Anderson 2008). Easton wrote in 1967 that "Possible fallouts from such a system are worldwide time synchronized to better than 0.1 microsecond" (Easton 1967). Thus, the time was more accurate by a factor of ten. In the Timation Development Plan (1971), the estimated time transfer accuracy was improved again by a factor of ten to better than 0.01 microsecond. Easton stated in 1974 that "precise orbiting clocks will prove to be a valuable tool in a variety of applications, by providing the entire planet earth with a single, accurate time system, enveloping the globe in a web of synchronized satellite signals" (Eberhart 1974). This was about 15 years before the term World Wide Web was coined. GPS and GNSS systems provide three-dimensional position and time. Even though time was less prominent than positioning at the time when GPS was formulated, it was extremely important for these pioneers of satellite navigation.

References

A Scientific Satellite Program, April 13, 1955, 10

R.E. Anderson, Early space age adventures. Quest: Hist. Spaceflight Q. **15**(3), 53 (2008a)

R.E. Anderson, A Navigation System Using Range Measurements with Cooperating Ground Stations, 317 (2008b)

R.L. Easton, Space Applications Branch Technical Memorandum No. 1. An Exploratory Development Program in Passive Satellite Navigation, May 8, 1967, 3

R.D. Easton, E.F. Frazier, *GPS Declassified: From Smart Bombs to Smartphones*, 42 (2013)

J. Eberhart, The Time Web. *Science News*, July 13, 1974, 27

M.P. Gleason, *Galileo: Power, Pride and Profit* (PhD diss., George Washington University, 2009), p. 5

D. Jewell Blogging on the GPS World website, January, 2008

C. McLaughlin, M. Lomask, *Project Vanguard: The NASA History* (Dover Publications, Inc., Mineola, NY, 2009), p. 190

G. Milner, *Pinpoint: How GPS is Changing Technology, Culture, and Our Minds* (W.W. Norton & Company, New York, 2016), p. 165

B.W. Parkinson, S.T. Powers, Part 2: The Origins of GPS, Fighting to Survive. *GPS World*, June 2010

Chapter 26
DASCH for Days to Decades Time Domain Astronomy

Jonathan Grindlay

Abstract The Digital Access to a Sky Century @ Harvard (DASCH) project has been underway for the past decade to digitize and fully reduce (photometry and astrometry) the ~450,000 glass plate images (not spectra) of the full sky taken by some ~30 Harvard telescopes from 1885 to 1992. The data processing pipeline achieves ~0.1 mag and ~0.5–3 arcsec (depending on plate series/scale) for all resolved stellar images (typically ~50,000) on each plate for the ~400 plates/day scanning and processing rate during full production. Nearly 1/3 of the plates are now online as of the 5th (of 12) data release, DR5. Unique time domain astronomy can be done for a vast range of objects; several examples will be shown. Of particular interest are rare extreme optical flares from accreting black holes, both stellar mass in black hole low-mass X-ray binaries (BH-LMXBs) and supermassive in active galactic nuclei (AGN). I will show examples of both. Just as full-production rate scanning/processing was to begin to finish by ~2018, a water main on Observatory Hill burst (on Jan 18, 2016) and flooded the DASCH lab, submerging some ~61,000 plates. These were quickly (within 2 days) removed to frozen storage (to prevent mold). Plate cleaning methods have now been developed (by the Harvard Weissman Preservation Center to fully restore them for scanning). Scanning is expected to resume in July 2016, with a new and faster scanner to enable the project to finish by late 2018.

Keywords Glass plate images • Scanning • Black holes • X-ray binaries • AGN

The Digital Access to a Sky Century @ Harvard (DASCH) project has been underway for the past decade to digitize and fully reduce (photometry and astrometry) the ~450,000 glass plate images (not spectra) of the full sky taken by some ~30 Harvard telescopes from 1885 to 1992. The data processing pipeline achieves ~0.1 mag and ~0.5–3 arcsec (depending on plate series/scale) for all resolved stellar images (typically ~50,000) on each plate for the ~400 plates/day scanning and processing rate during full production. Nearly 1/3 of the plates are now online as of

J. Grindlay (✉)
Harvard University - CFA, Cambridge, USA
e-mail: jgrindlay@cfa.harvard.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_26

the 5th (of 12) data release, DR5. Unique time domain astronomy can be done for a vast range of objects; several examples will be shown. Of particular interest are rare extreme optical flares from accreting black holes, both stellar mass in black hole low-mass X-ray binaries (BH-LMXBs) and supermassive in active galactic nuclei (AGN). I will show examples of both. Just as full-production rate scanning/processing was to begin to finish by ~2018, a water main on Observatory Hill burst (on Jan 18, 2016) and flooded the DASCH lab, submerging some ~61,000 plates. These were quickly (within 2 days) removed to frozen storage (to prevent mold). Plate cleaning methods have now been developed (by the Harvard Weissman Preservation Center to fully restore them for scanning). Scanning is expected to resume in July 2016, with a new and faster scanner to enable the project to finish by late 2018.

Chapter 27
Mean Solar Time and Its Connection to Universal Time

John H. Seago and P. Kenneth Seidelmann

Abstract Universal Time is the measure of Earth rotation that also serves as the astronomical basis of civil timekeeping. As the successor to Greenwich Mean Time (GMT), Universal Time intended to maintain uniform time as the angle between zero longitude and the mean sun—a fictitious point moving uniformly along the celestial equator that keeps pace with the Sun over the very long term. However, variability in the rotation rate of the Earth has caused Universal Time to diverge from the original geometric concept of mean solar time by approximately (1/365.2422) ΔT, where ΔT is the accumulated measure of nonuniform Earth rotation since 1900. After accounting for changes in the origin of the terrestrial and celestial reference systems since the end of the nineteenth century, simulated transits confirm that the Sun on average now crosses zero longitude at 12 h 00 m 00.2 s Universal Time, a result that is expected from theory.

Keywords Equation of time • Greenwich mean time • Ephemeris time • Mean solar time

Introduction

The duration of the solar day is the time interval between two culminations of the apparent Sun over a meridian, with *apparent solar time* being the hour angle of the Sun between these culminations. Apparent solar days have irregular duration, due to a nonuniform motion of the Sun in right ascension, caused by the obliquity of the ecliptic and the eccentricity of the Earth's orbit. The irregularity of the Sun's motion is often illustrated through graphs of the *equation of time*, the difference between apparent solar time and mean solar time. *Mean solar time* provides uniform hours

J.H. Seago (✉)
Analytical Graphics, Inc., 220 Valley Creek Boulevard, Exton, PA 19341-2380, USA
e-mail: jseago@agi.com

P.K. Seidelmann
Astronomy Department, University of Virginia, 400325
Charlottesville, VA 22904, USA
e-mail: pks6n@virginia.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_27

and days that keep pace with the Sun on average over the long term and is therefore amendable to timekeeping with clocks (Newcomb and Holden 1880). In the same way that successive culminations of the Sun on the celestial sphere define the apparent solar day, the mean solar day implies sequential culminations of a fictitious point moving along the celestial equator, historically known as the *mean* sun; mean time of day is defined by the hour angle of this fictitious sun (Woodhouse 1812, pp. 222–224). As a mathematical abstraction traveling at uniform angular velocity, this mean sun cannot be observed directly.

Mean Solar Time

The concepts underlying mean solar timekeeping have been known since antiquity. Aristotelian astronomy held that heavenly bodies traveled according to uniform, circular motion or else from a combination of uniform motions. The seasonal irregularity of the Sun is recognized as a deviation from the uniform model; the relationship between apparent solar time and average solar time was empirically determined as early as the second century A.D. by Ptolemy from lunar studies (Neugebauer 1975). The European medieval timekeeping device known as the *horologium nocturnum* (or nocurlabe or "nocturnal") was built upon the basic principles of mean solar timekeeping and may have been described in Chinese annals as an astronomical "star dial" as early as *c*.1280 (Sivin 2009). But the proliferation of well-regulated mechanical clocks into the eighteenth and nineteenth centuries ushered mean solar timekeeping into general usage, to the point where "mean time" became synonymous with time "o'clock," to differentiate mean time from sundial time.

The term "fictitious mean sun" thus came into astronomical usage, referring to a convenient value of right ascension that travels uniformly along the celestial equator, with an origin and rate equal to that of the Sun's mean longitude along the ecliptic. The term (in English) has enjoyed precise usage since the eighteenth century at least. Before the first publication of the British *Nautical Almanac and Astronomical Ephemeris* for celestial navigation, Maskelyne (1764) described mean time as:

> . . .what would be pointed out by the ſun, if his motion in right aſcenſion from day to day was uniform, or, in other words, it is what would be pointed out by a fictitious ſun or planet ſuppoſed to move uniformly in the equator, with a motion equal to the mean motion of the ſun in longitude, its diſtance from the firſt point of Aries (meaning hereby the mean equinox) being always equal to the mean longitude of the ſun: and as apparent noon in the inſtant of the true ſun's coming to the meridian, ſo mean noon is the inſtant at which this fictitious planet would come to the meridian. The interval between its coming to the meridian on any two ſucceſſive days is a mean ſolar day, which is divided into hours, minutes, & c. of mean ſolar time; all which it is manifeſt will preſerve the ſame length at all times of the year.

Mean solar time on the astronomical meridian of the Greenwich observatory, or GMT, was recommended as a global standard by 1884, although disparate

astronomical theories for reckoning mean solar time were in use at that time. Recognizing the problem, Simon Newcomb, director of the Nautical Almanac Office of the US Naval Observatory from 1877 to 1897, strove to develop a self-consistent set of standard astronomical constants (Newcomb 1883). Newcomb's constants served to further unify the calculation of astronomical ephemerides and phenomena internationally, and his 1895 expression for the right ascension of the mean sun remained operationally unchanged until 1984. Subsequently, over the twentieth century, the term "fictitious mean sun" became synonymous with Newcomb's particular realization (Green 1985), i.e.,

$$R_{\odot}(T) = 18\ \mathrm{h}38\ \mathrm{m}45.836\ \mathrm{s} + 86401845.42\ \mathrm{s}T + 9.29\ \mathrm{s}T^2, \tag{27.1}$$

where T is the number of Julian millennia of 365,250 days elapsed since "1900, Jan. 0, Greenwich Mean noon."[1] Mean solar time at Greenwich, measured from midnight, was expressible as the hour angle of this fictitious mean sun +12 h or equivalently (Explanatory Supplement 1961, p. 74)

$$\begin{aligned}\mathrm{GMT} = {} & 12\ \mathrm{h} + \text{Greenwich hour angle of the mean equinox of date} \\ & - R_{\odot}(T).\end{aligned} \tag{27.2}$$

Thus, mean solar noon (12 h) occurred at Greenwich whenever the Greenwich hour angle of the mean equinox of date equaled $R_{\odot}(T)$. The academic significance of the fictitious mean sun, having right ascension $R_{\odot}(T)$, was that its transit over any celestial meridian defined the moment of mean noon for that meridian.

Newcomb did not specify the fundamental measure of time that T reckoned in $R_{\odot}(T)$, although basic units of mean solar days were always presumed (Seidelmann 1992). In Newcomb's era, no distinction was cast between the progression of time indicated by the independent variable of a solar system theory and the progression of time of day indicated by the rotation of the Earth. These two concepts were not fully separated until the mid-twentieth century, after the Earth's rotation was concluded to be slightly nonuniform. Subsequently, the theoretically uniform argument of Newcomb's solar system theory became known as *Ephemeris Time* (ET). In principle, Ephemeris Time could be determined by observing the locations of solar system objects and was intended for scientific applications. Meanwhile, the measure of the Earth's rotation persisted as the basis of civil timekeeping and calendrical dating, under the separate name *Universal Time* (UT).

Accordingly, Newcomb's T of Eq. (27.1) was eventually interpreted as the independent variable of the Earth's orbital motion and became associated with theoretically uniform Ephemeris Time (T_{E}). This implied the existence of an "ephemeris mean sun" $R_{\odot}(T_{\mathrm{E}})$ with right ascension equal to Newcomb's fictitious mean sun. Meanwhile, the hour angle of Earth's zero longitude continued to be

[1] Or 31 December 1899, 12 h UT. Newcomb originally used units of Julian centuries instead of Julian millennia; a larger unit is adopted here to be consistent with modern expressions.

measured relative to the equinox, and Universal Time (T_U) was employed within Eq. (27.1) to effectively define a "universal mean sun," $R_\odot(T_U)$. $R_\odot(T_U)$ thus became a conventional expression that related Universal Time and Greenwich mean sidereal time (GMST) at 12 h UT, per Eq. (27.2). Because T_U was used in Eq. (27.1), rather than T_E, GMST was not a precise function of the right ascension of Newcomb's mean sun, so Universal Time was not rigorously defined as the hour angle of Newcomb's fictitious mean sun increased by 12 h. Instead, UT was practically defined by operational procedures employing the conventional relationship of Eq. (27.2), with T_U as the independent variable of $R_\odot(T)$. The difference between $R_\odot(T_E)$ and $R_\odot(T_U)$ was expressed as (Explanatory Supplement 1961, p. 78)

$$R_\odot(T_E) - R_\odot(T_U) \approx 0.002738\,\Delta T, \tag{27.3}$$

where ΔT is the accumulated excess of the measure of Ephemeris Time over Universal Time, ET − UT.[2]

Universal Time

Continuing operational improvements led to refined versions of Universal Time (UT0, UT1, UT2) having varying degrees of uniformity with periodic differences at the level of tens of milliseconds in the 1950s. A broadcast convention for Universal Time, related to atomic frequency standards, became known as *Coordinated Universal Time* (UTC) in the 1960s. By the 1980s, traditional optical astrometry had been replaced by new technologies that offered better accuracy in the measurement of Earth rotation. By this point, only two versions of Universal Time found widespread application: UT1 as a precise measure of Earth's terrestrial origin (prime meridian) about its observed rotational pole and UTC as the atomic realization of Universal Time broadcast for precision work and civil timekeeping. UT1 is made available to high accuracy by adding a correction to UTC:

$$\mathrm{UT1} = \mathrm{UTC} + (\mathrm{UT1} - \mathrm{UTC}) = \mathrm{UTC} + \Delta\mathrm{UT1}. \tag{27.4}$$

ΔUT1 is tabulated by the International Earth Rotation and Reference Systems Service (IERS), a service chartered in 1988 by the International Astronomical Union (IAU) and the International Union of Geodesy and Geophysics to supplant the *Bureau International de l'Heure* (BIH) and coordinate the results of astro-geodetic observing programs.

[2] After 1978, ΔT came to represent the accumulated difference between Universal Time and the theoretically uniform time scale Terrestrial Time—the modern successor to Ephemeris Time within the framework of general relativity.

Today, UT1 is determined fundamentally from observations of extragalactic radio sources using Very-Long-Baseline Interferometry (VLBI). Prior to 2004, Universal Time was defined in terms of mean sidereal time, but it has since been defined in terms of its relationship to the *Earth rotation angle* (θ), the angle of the Terrestrial Intermediate Origin from the Celestial Intermediate Origin (Capitaine et al. 2000):

$$\theta = 67310.54841001 + 86636.546949141027086\,T_u. \tag{27.5}$$

where T_u is the number of days of UT1 elapsed since Julian Date 2451545.0 UT1. Under this definition, the time derivative of UT1 is directly proportional to the angular rotation rate of Earth. The constant of proportionality between θ and T_u was chosen to maintain continuity with previous definitions of Universal Time and, therefore, has traceability back to Newcomb's expression for the right ascension of the mean sun. Consequently, UT1 is still considered to be nominally equivalent to mean solar time reckoned from Greenwich (Urban and Seidelmann 2012).

A debate began *c*.1998 as to whether Coordinated Universal Time should remain subject to the occasional intercalary adjustments known as *leap seconds*, which keep UTC within 0.9 s of Universal Time (Seago et al. 2011, 2013). One of the advertised benefits of leap seconds is to synchronize atomic time with the Sun, but the relationship between Universal Time and the geometry of the Sun is no longer obvious. The historical development of mean solar time prior to the realization of variability in Earth rotation, and the implications of this discovery to Universal Time, is described heuristically in the Explanatory Supplement (1961) and subsequent editions, but the motivations for such discussions are likely to be understood only by audiences already familiar with the underlying technical issues. The distinction between the measure of Universal Time and the geometry of mean solar time has since become historically opaque, as modern textbooks deemphasize the role of Universal Time as an indicator of mean solar time at Greenwich, understanding that the separation has been small. Contemporary estimates of the slight separation between the operational realization of Universal Time and geometric mean time have varied from "1 second over a few centuries" (Audoin and Guinot 2001) to "a departure which may reach 1 or 2 seconds," depending on the effect being described (Guinot 2011). Reasons for the separation include:

- The timing of solar transits cannot be precisely measured with respect to the background stars, such that historical determinations of UT with respect to the Sun may be biased on the order of 0.1 s.
- Universal Time was estimated by the BIH using a "mean observatory," determined through a weighted average of Universal Times from radio signals of contributing observatories relative to their adopted longitudes (*heure definitive*). Changes to the averaging algorithms over the years may have preserved continuity at the order tens of milliseconds.
- Plate tectonics cause the relative longitudes of stations to change on the order of centimeters (microseconds) per annum.

- The effects of polar motion cause declination-dependent nonuniformities in the time observed (on the order of 30 ms at Greenwich latitude), requiring the IAU to place the terrestrial fiducial direction for Universal Time at zero latitude.
- A change of terrestrial conventions from astronomical to geodetic coordinates places the IERS zero meridian 0.089′ (5.34″) east of the Airy Transit Circle, the traditional landmark indicating zero longitude. However, the difference between astronomical to geodetic coordinates—deflection of the vertical—is a local phenomenon that does not present a systemic change in the global origin for UT1 (Malys et al. 2015).

A detailed error budget of these combined effects over time would be very difficult to construct, and there may be some debate as to what the phrase "mean solar time at Greenwich" implies in any analysis of the question. But this situation does not prevent a comparison of the IERS Universal Time series with a specific mathematical model for the motion of a "mean sun." Newcomb's expression for mean solar time, Eq. (27.1), has the advantage of being the only historical convention internationally adopted for such purposes. The comparison of UT1 with an expression based on more modern solar system theory also presents an opportunity to assess the question: how well does Universal Time keep pace with the actual Sun over the very long term?

The Geometry of Mean Solar Time

The duration of the mean solar day may be reckoned, most simply, by counting the number of apparent solar days over the course of a solar year, i.e.,

$$d_{\odot} = Y/k_{\odot}, \tag{27.6}$$

where $d_{\odot}$ represents the duration of a *mean solar day*, Y represents the duration of one solar year, and $k_{\odot}$ is a fractional number of Earth rotations relative to the Sun over Y. Since antiquity, this estimate was made more accurate by averaging over multiple solar years, where a solar year was measured as the time required by the Sun to reach a particular declination or ecliptic longitude. Based on the number of days between solstice observations centuries apart, Hipparchus estimated that the duration of the solar year was several minutes shorter than 365 ¼ days, the customary value of his era. Measurement of the annual period between the solstices originally led to the term *tropical year* (*tropics* meaning "turn"), when the declination of the Sun changed direction (Woodhouse 1812, pp. 64–66). This term apparently evolved to also describe the duration between annual solar passages at the equator, or vernal equinox (Gummere 1842); this *vernal equinox year* was critical to establishing the Gregorian calendar for determining the date of Easter (Steel 2000).

Successive annual passages of the solstices or equinoxes are not strictly uniform, because the orbital motion of the Earth's line of apses relative to the equinox introduces tiny rate differences, which differ depending on the starting location within the orbit (Meeus and Savoie 1992). The intersection of the planes of the equator and ecliptic, which traditionally defines the equinox direction, also has secular motion relative to inertial space due to precession and periodic nutation of the orientation of the Earth's spin axis and equator, primarily caused by gravitational torqueing of the Earth's figure by the Moon. Due to natural perturbations to the Earth's orbit and depending on the frames of reference used, variations in the time between successive annual passages of the Sun can vary by several minutes. These variations, and the advantage of averaging observations over multiple annual cycles, are illustrated in Fig. 27.1.

Tropical Year

Today, a *tropical year* is most accurately defined as the period by which it takes the mean orbital longitude of Earth to advance 360° (one revolution or cycle) on average. This modern definition results from the practice of relating the nonuniform motion of the Earth-Moon barycenter to the independent argument of a precise solar system theory. Specifically, if the mean longitude of the Sun, L, can be generally expressed in the time polynomial form (relative to a precessing mean equinox of date)

$$L = L_0 + L_1 t + L_2 t^2 + L_3 t^3 + \cdots, \tag{27.7}$$

then the (precessing) mean motion can be expressed as the time rate of change of L:

$$dL/dt = L_1 + 2L_2 t + 3L_3 t^2 + \cdots. \tag{27.8}$$

To express the duration of the tropical year, Y, in units of Barycentric Dynamical Time (TDB) seconds, one takes the reciprocal of Eq. (27.8) and applies appropriate scale factors to convert from units of arcseconds and Julian millennia (Smith et al. 1989):

$$Y = \frac{(360 \times 60 \times 60) \times 325250 \times 86400}{L_1 + 2L_2 t + 3L_3 t^2 \quad + \cdots} \tag{27.9}$$

If $L_1 >> L_2$, L_3, etc., then Eq. (27.9) can be conveniently expressed as (Borkowski 1991)

$$Y = \frac{(360 \times 60 \times 60) \times 365250 \times 86400}{L_1} \left(1 - 2\frac{L_2}{L_1} T - 3\frac{L_3}{L_1} T^2 - \cdots\right). \tag{27.10}$$

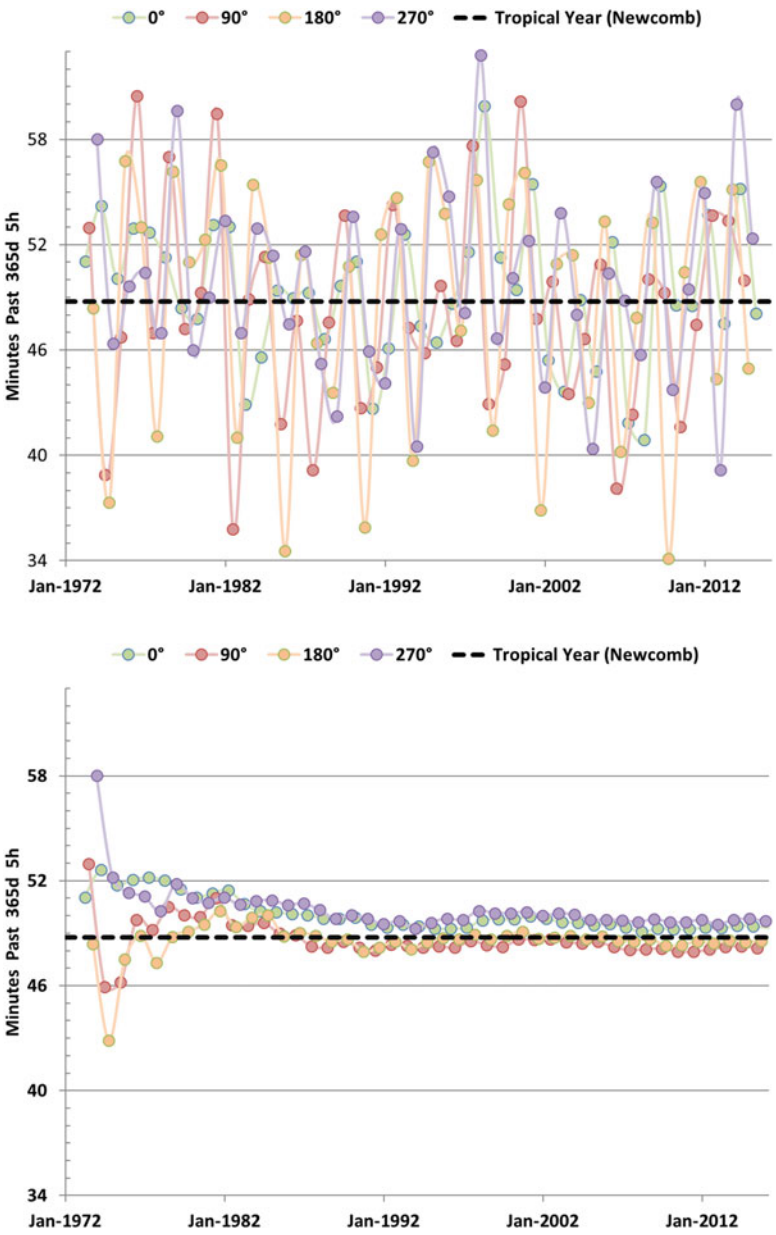


Fig. 27.1 *Top*: durations (in SI units) between successive solar longitude crossings from 1972 to 2015. *Bottom*: averages of the same series, reestimated each year by including an annual observation. The averages approach Newcomb's tropical year (*dotted line*) after several decades. Averaging the duration between seasonal events over longer intervals significantly improves the accuracy of the estimate, but careful inspection reveals that the annual averages for various solar longitude crossings differ by about a minute

where T is in units of Julian millennia. Using Eq. (27.1) as an example, Newcomb's tropical year is 31556925.9747 s–5.3032 s T from epoch 1900.

Sidereal Time

An expression similar to Eq. (27.6) may be developed for the mean *sidereal day*, $d_\bigstar$, as the duration of the Earth's rotation between two successive culminations of the precessing vernal equinox (the point where the ascending sun notionally crosses the celestial equator):

$$d_\bigstar = Y/k_\bigstar. \tag{27.11}$$

Because the Earth circuits the Sun exactly once during P, the Earth has exactly one less rotation relative to the Sun than to the stars, such that $k_\bigstar = k_\odot + 1$ (Newcomb 1906). This allows $k_\bigstar$ and $k_\odot$ to be entirely eliminated when combining Eqs. (27.6) and (27.11), to obtain an expression of frequency $(d_\odot)^{-1} = (d)^{-1}$–$(Y)^{-1}$ or similarly

$$\Omega_{\mathrm{MT}} = \omega - n. \tag{27.12}$$

Here, Ω_{MT} is the rate of *mean time of day*, ω is the rate of Earth rotation with respect to the equinox (*sidereal time*), and n is the rate of solar *mean motion*.

According to Eq. (27.12), the geometrical interpretation of mean time of day is the dihedral angle of the celestial meridian moving at an angular sidereal rate ω, relative to the Sun's mean motion moving at rate n with respect to the mean equinox. One can visualize the combination of these two distinct rotations as the Earth spinning in space, with respect to a mean time "dial" that slowly rotates in the plane of the celestial equator at a rate of once per year (Fig. 27.2). The orientation of the imaginary celestial dial is chosen so that 12 h corresponds to local noon at zero longitude, thus making the 12 h mark of this imaginary dial the "mean sun." Local time always corresponds to the value of the imaginary equatorial dial transiting the local meridian, or line of longitude.

Because the mean sun cannot be measured directly, mean solar time was practically obtained by multiplying the measure of sidereal time by a conventional proportion, e.g.,

$$\Omega_{\mathrm{UT}} = \omega\, S. \tag{27.13}$$

To relate this to the geometrical concept of Ω_{MT} in Eq. (27.12), S must be $(1 - n_\oplus/\omega)$ or, equivalently, $(k - 1)/k$. A value of $k/(k - 1) = 1.002737909265 + 0.589 \times 10^{-10}\, T_{\mathrm{U}}$ (T_{U} being the number of UT days since noon, Dec. 31, 1899) arises from Newcomb's solar system theory; however, the secular variation of one-half second per century was considered "inappreciable"

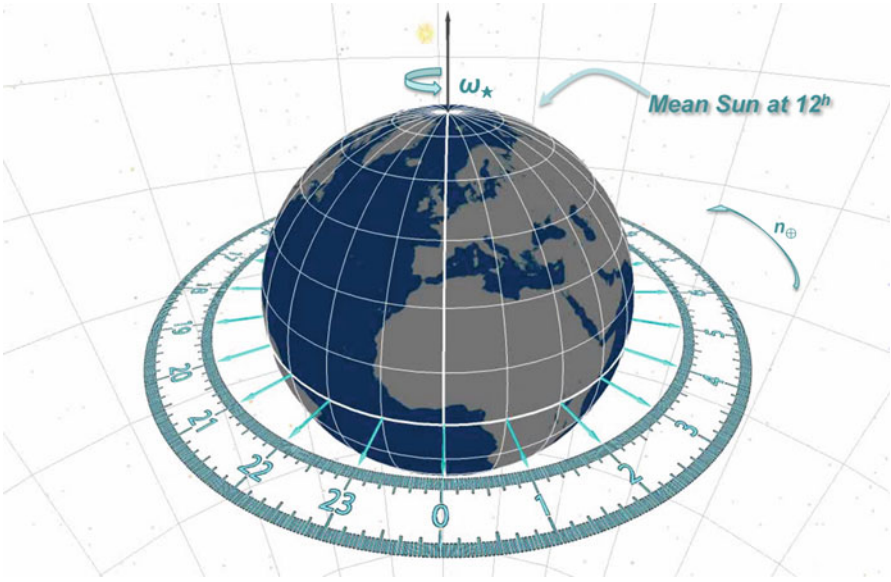


Fig. 27.2 Earth as a clock (0 h UT, February 29, 2000). The Earth spins at rate $\omega_\star$, while the mean time "dial" rotates once per year with the fictitious mean sun at 12 h. Each meridian indicates its own mean solar time

when relating the proportion of mean solar time to mean sidereal time (Explanatory Supplement 1961, p. 76).

The traditional use of a scale factor implied that S and, therefore, ω and k were fixed. Newcomb (1906) was "obliged, in all ordinary cases, to treat [Earth rotation] as invariable" despite there being "theoretical reasons for believing that the speed of rotation is slowly diminishing from age to age." Constancy of Earth rotation was the prevailing assumption until the early twentieth century: in 1675 Flamsteed tested this hypothesis using the best chronometers of his era and concluded that Earth's rotation was constant (Howse 1997). Laplace also concluded in 1799 that the inequalities of Earth rotation rate were insensible; likewise, Serret in 1859 inferred that the period of rotation must be practically constant due to the lack of detectable polar motion at the time (Ball 1880). However, by the mid-nineteenth century, Delaunay had opined that a change in the rate of Earth's rotation should be sensible from ancient astronomical records (Ibid.); for example, Schjellerup (1881) found that a quadratic correction was required to represent very ancient Chinese eclipses.

Because Earth's rotation ω is less than uniform, an integration of Eq. (27.12) will slowly separate from an integration of Eq. (27.13) by an amount that is proportional to ΔT. To show this, first, an angular analogue to uniform Ephemeris Time is taken from Eqs. (27.12) and (27.14) as

$$\Omega_{\mathrm{ET}} = \varpi - n = \varpi, \widehat{S}, \tag{27.14}$$

which assumes a perfectly constant sidereal rate of ϖ, with fixed $\hat{S} = (1 - n_{\oplus}/\varpi)$. When an invariable scale factor $\hat{S}$ is adopted for use in Eq. (27.13), the difference between Eqs. (27.12) and (27.13) becomes

$$\Omega_{\mathrm{MT}} - \Omega_{\mathrm{UT}} = \Delta\omega \left(1 - \widehat{S}\right) = \Delta\omega \, n/\varpi = (\Omega_{\mathrm{UT}} - \Omega_{\mathrm{ET}})n/\Omega_{\mathrm{ET}} \tag{27.15}$$

where $\omega = \varpi + \Delta\omega$. Recognizing that $\int(\Omega_{\mathrm{UT}} - \Omega_{\mathrm{ET}})\, dt = \Delta T$, the integral of Eq. (27.15) is

$$\int \Omega_{\mathrm{MT}}\, dt - \int \Omega_{\mathrm{UT}}\, dt = \Delta T/(\Omega_{\mathrm{ET}}/n\,) \tag{27.16}$$

where $(\Omega_{\mathrm{ET}}/n_{\oplus})$ is the number of ephemeris days in the tropical year, or $1/(k-1)$. Consequently, after converting angular cycles to the unit of days, there is a theoretical expectation that

$$(t_{\mathrm{MT}} - t_{\mathrm{UT}}) \approx \Delta T/365.2421987768 \tag{27.17}$$

where t_{MT} is mean time implied by geometry from the mean sun and t_{UT} is astronomical time implied by rescaling the rotation rate of the Earth. The Explanatory Supplement (1961) provides the result of Eq. (27.17) as 0.002738 ΔT, within its development of the so-called ephemeris meridian, an auxiliary reference meridian conceived to locate the longitude of origin supposing that Earth rotation had been uniform. Operational use of the ephemeris meridian is almost a curiosity today, except for finding use in the calculation of apparent solar time via the "Equation of Ephemeris Time." Because the SI second was first defined in terms of Ephemeris Time and the atomic second was calibrated to that measure, the duration of the atomic SI second is based on Newcomb's mean sun, but precise connections are now rarely drawn between Ephemeris Time and the geometry of mean sun.

Mean Solar Time and the Equation of Time

Universal Time, as it is presently employed, has separated from the original geometric concept of mean solar time at Greenwich by about 0.2 s and will continue to separate as a function of ΔT. This conclusion is not only supported by academic considerations, such as first principles according to Eq. (27.17) , but it is supported by precise simulations of the fictitious mean sun as an Earth-orbiting object (Seago and Seidelmann 2013). This result can also be shown from computed transits of the apparent Sun at zero longitude, corrected for the equation of time.

The equation of time is fundamentally defined as the hour angle of apparent Sun minus the hour angle of the mean sun, the 12 h point identified in Fig. 27.2

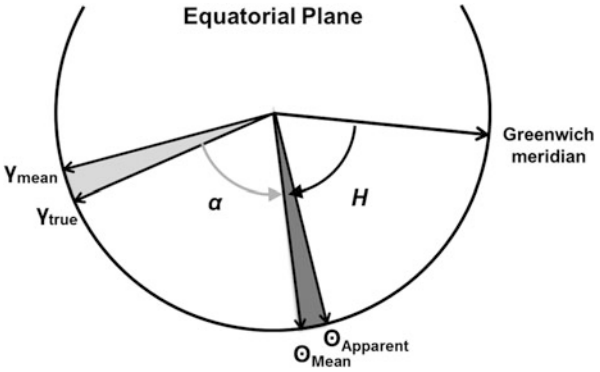


Fig. 27.3 The *equation of time* is the Greenwich Hour Angle (H) of apparent Sun minus the GHA of mean sun or the right ascension (α) of the mean sun minus α of the apparent Sun, α referring to the true equinox (γ_{true}). The *equation of the equinoxes* is the α of the mean equinox (γ_{mean})

(Hughes et al. 1989). Hour angles may be defined with respect to any local meridian, although "Greenwich" (zero longitude) is traditionally employed as the reference meridian due to convenience:

$$\text{equation of time} = \text{GHA}_{\text{Apparent}\odot} - \text{GHA}_{\text{Mean}\odot}, \tag{27.18}$$

When referring to Greenwich, an hour angle may be thought of as the terrestrial longitude over which an astronomical object is positioned at that moment.

From Fig. 27.3, it is also evident that the equation of time may be expressed as the difference in right ascensions, instead of the difference in hour angles or longitudes. This leads to an alternative formulation sometimes found in textbooks:

$$\text{equation of time} = \alpha_{\text{Mean}\odot} - \alpha_{\text{Apparent}\odot}, \tag{27.19}$$

The equation of time is illustrated in Fig. 27.4 over one annual period. The departure of the Sun from its mean never exceeds about 16 min, or a change of about 1% along the equator.

To create a uniform scale from solar time, Newcomb defined his fictitious mean sun to have strict uniform motion μ along the celestial equator in the reference frame defined by the best precessional constants of the 1890s, with a right ascension α as nearly as possible to the Sun's mean longitude L, so that $\alpha \approx L$. The total rate of the mean sun along the equator approximates the average rate of the Sun's ecliptic longitude, plus the average rate of the equinoctial precession along the ecliptic. This total rate amounts to one cycle per tropical year, according to the modern definition of Y described in Eq. (27.9). Meanwhile, the mean longitude of the apparent Sun, referred to a precessing mean equinox of date and displaced by aberration, is expressed in the time polynomial form as (Woolard and Clemence 1966)

$$L = \left(\lambda_0 + \lambda_1 t + \lambda_2 t^2 + \cdots\right) + \left(h_1 t + h_2 t^2 + \cdots\right) - \kappa. \tag{27.20}$$

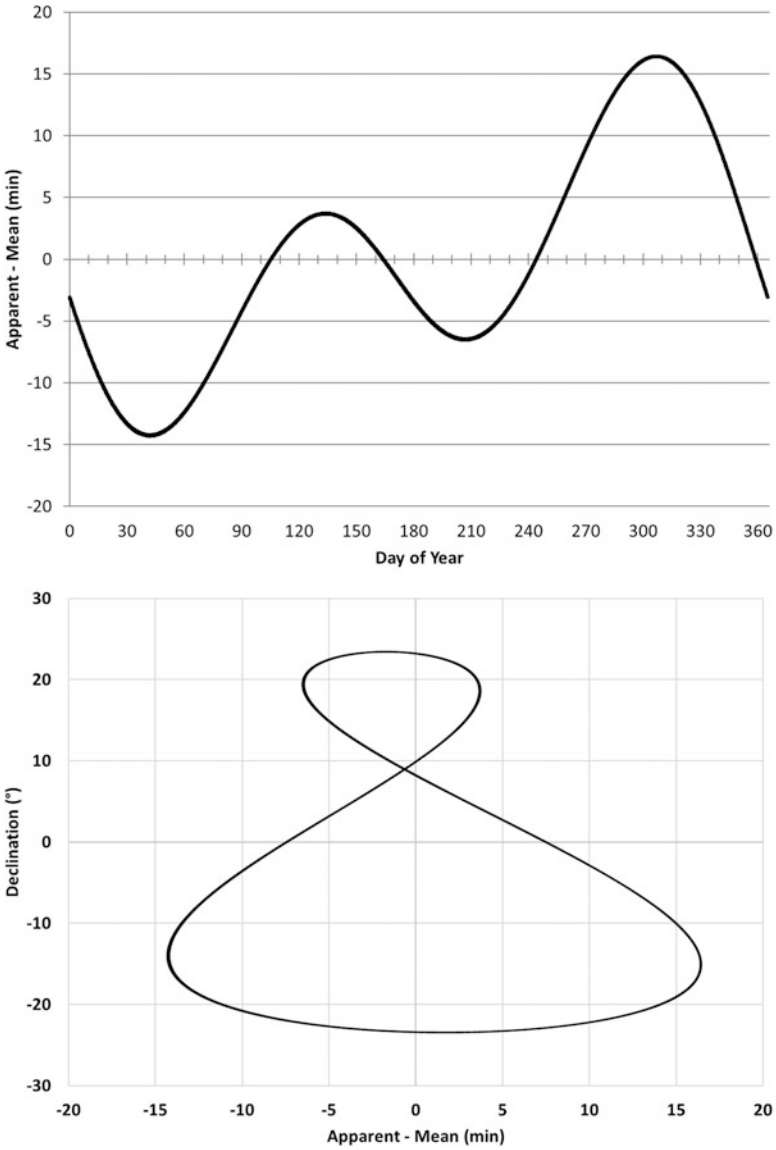


Fig. 27.4 The equation of time, as a function of time of year (*top*) and solar declination (*bottom*)

Here, terms of λ_i represent the ecliptic mean motion of the apparent Sun, terms of h_i represent general precession in longitude (along the ecliptic), and κ represents the constant of aberration. Comparing Eq. (27.20) with Eq. (27.7), $L_0 = (\lambda_0 - \kappa)$, and $L_i = (\lambda_i + h_i)$; $i > 0$. Meanwhile, the expression for the right ascension of the mean sun can be formulated as

$$L \approx \alpha = \alpha_0 + \mu t + \left(m_1 t + m_2 t^2 + \cdots\right). \tag{27.21}$$

Here, terms of m_i express the motion of general precession of the equinox in right ascension (along the equator) as the reference point for right ascension α. Equating the coefficients of like terms between Eqs. (27.7) and (27.21), one notices that $\alpha_0 = L_0$ and $L_1 = (\lambda_1 + h_1) = (\mu + m_1)$. Therefore, the right ascension of the mean sun is expressible as

$$\alpha = L_0 + L_1 t + \left(m_2 t^2 + \cdots\right) = (\lambda_0 - \kappa) + (\lambda_1 + h_1)t + \left(m_2 t^2 + \cdots\right). \tag{27.22}$$

Although Maskelyne and others prescribed the right ascension of the mean sun "as equal" to the Sun's aberrated mean longitude along the ecliptic, Newcomb amended this prescription to be "as nearly as may be" possible, in order to make mean solar time more theoretically uniform. Newcomb noted that the nonlinear, secular accelerations expressed within the ecliptic (terms of $\lambda_2 t^2 + h_2\ t^2 + \ldots$) would not equal the nonlinear accelerations expressed in the equator (terms of $m_2 t^2 + \ldots$); therefore, a linear representation of the motion of the mean sun in the equator would necessarily diverge from the mean longitude of the actual Sun. The differences in these higher-order terms are negligible only when t is relatively small—near the origin for time. Newcomb estimated the amount of separation, or "error," to be 2.0 s T^2 (2 s per millennium squared or about 0.02 s by the end of the twentieth century), the correction of which "we must leave to astronomers of the future" (Newcomb 1895a). de Sitter (1938) noted that Newcomb's fictitious mean sun would need to be altered so that it is *exactly* equal to the Sun's mean longitude, resulting in equatorial motion that is not strictly uniform. The operational practice of the nineteenth century was to improve the system of constants and solar system theory every few decades, due to their uncertainty; the origin of T was occasionally relocated to a new epoch, and higher-order terms of T^2 remained negligibly small.

Some examples of changes to the operational values of the fictitious mean sun, as adopted by nineteenth-century astronomical almanacs, are expressed in Table 27.1 (Stone et al. 1883). To make the values comparable, the expressions are referred to a common epoch of January 1, 1850, Paris, mean noon. A notable difference exists between the expressions of Bessel and Hansen; Hansen discovered a previously un-modeled long-period inequality[3] whereby Jupiter and Mars affect the position of Earth by almost 1 s in time over eighteen centuries. Because this inequality was approaching its minimum by 1850, its inclusion in Hansen's solar theory caused his mean ellipse for the Earth's orbit to change noticeably relative to previous theories, thereby altering the origin and rate of the fictitious mean sun based on that ellipse. Newcomb questioned such noticeable alteration of mean sun

[3]The term "inequality" describes the tendency of a planet to stray from an elliptical path. Low-order series solutions of planetary equations via perturbation theory do not model the long-period inequalities of near-resonant periodic behavior. The accounting of long-period inequalities offered significant improvements for nineteenth-century ephemerides.

Table 27.1 Nineteenth-century values for the origin and rate of the fictitious mean sun, referred to a common epoch of January 1, 1850, Paris, mean noon

Source	α_0	$\dot{\alpha}$ (seconds/Julian millennium)
Bessel (*c*.1828)	18 h 43 m 6.41 s	86401841.212
Hansen (*c*.1853)	18 h 43 m 6.88 s	86401844.937
LeVerrier (*c*.1859)	18 h 43 m 6.90 s	86401845.227
Newcomb (1895a)	18 h 43 m 6.90 s	86401844.490
Newcomb (1898)	18 h 43 m 6.89 s	86401844.491

The mean longitude and mean motion depend on the adopted solar system theory. Values reflect a mean ellipse, best fit to the Earth's orbit

adopted by almanacs, because it placed the mean sun away from the true Sun by more than one-third second of time in his era and presented a noticeable discontinuity with earlier solar tables. Interestingly, between the publication of Newcomb (1895a) and Newcomb (1898), Newcomb could not resist tuning his own expression for the mean Sun (Table 27.1), following his own investigation of long-term inequalities (Newcomb 1895b).

UT1: Transit Analysis of the Mean Sun and Apparent Sun

Although the definition of Universal Time has changed since Newcomb's era, changes are reputed to have maintained continuity with Newcomb's adopted constants. Seago and Seidelmann (2013) simulated Newcomb's fictitious mean sun over four decades as a circular Earth satellite and calculated the Universal Times when the direction to the geometric mean sun was parallel to the ITRF zero meridian plane. The experiment demonstrated that the degree of separation between Newcomb's mean sun and UT noon was compatible with theoretical expectations.

An alternative approach was used for this paper, namely, to model transits of the apparent Sun, instead of the mean sun, according to a solar ephemeris. This approach is equivalent to modeling transits of the mean sun (see Fig. 27.3), except that the Universal Times of solar transits should equal the equation of time as defined by Eq. (27.19). To use Eq. (27.19), α must be expressed relative to the true equinox of date, a subtle requirement often neglected by modern treatments of the subject. This subtlety is noteworthy because expressions for the right ascension of the mean sun, such as Eq. (27.1), are with respect to the *mean* equinox of date. The difference amounts to nutation in right ascension, known as the *equation of the equinoxes*, which, if neglected, can cause a declination-dependent error as large as a few tenths of a second in the determination of the transit time.

The product of the simulation was, therefore, the equation of time at the moment of simulated transit per Eq. (27.19), minus the value of UT1 at the moment of simulated transit. For the transit simulation, the JPL Developmental Ephemeris 421 (DE 421) was used to model the relative position between the Earth and Sun

within the *Systems Tool Kit* (STK) software, a commercial software product by Analytical Graphics, Inc. (AGI), available to educational institutions without cost. Each transit time was modeled as the moment at which the dihedral angle became zero between the direction of the apparent Sun (affected by aberration) from the geocenter and the zero meridian plane. That plane includes the Celestial Intermediate Pole as defined by IERS Conventions and the ITRF X-axis as the longitude origin. The equation of time was modeled according to Eq. (27.19), with the right ascension of the apparent Sun, $\alpha_{\text{Apparent}\odot}$, calculated relative the true equinox as defined by the IAU 1976 precession and 1984 nutation theories.

The right ascension of Newcomb's mean sun, $\alpha_{\text{Mean}\odot}$, was modeled in Eq. (27.19) with respect to the mean equinox of J2000 as

$$R_{\odot} = 1009658.2260'' + 1296027721.93''T_{\text{E}} + 139.4''T_{\text{E}}{}^2[\text{epoch}\,2000]. \quad (27.23)$$

where T_{E} is in units of Julian millennia TDB from epoch J2000.0. The development of Eq. (27.23) is detailed by Seago and Seidelmann (2013); it differs from Eq. (27.1) primarily due to an update of the origin from epoch 1900 and an FK4 equinox adjustment (Kaplan 1981). To refer $\alpha_{\text{Mean}\odot}$ to the true equinox used by STK, the equation of the equinoxes (based on IAU 1984 nutation) was applied to Eq. (27.23) before $\alpha_{\text{Apparent}\odot}$ was subtracted.

The results of the simulation are illustrated with Fig. 27.5. The simulated transit times of the apparent Sun across the ITRF zero meridian plane, adjusted for the equation of time based on Newcomb's expression for right ascension of the Fictitious Mean Sun as defined by Eq. (27.19), occur a few tenths of a second after noon UT1. The trend is proportional to ΔT, in accordance with Eq. (27.17). If Eq. (27.17) is subtracted from the simulated transit times, the result should be zero (within the precision of the calculation). This is illustrated in Fig. 27.6, where a uniform spread in the residuals is evident due to numerical cutoff in the iterated search for the simulated UT1 transit times. These searches did not seek to preserve accuracy in UT1 below ± 0.5 ms, understanding that Newcomb's original expression for the mean sun was only precise to 1 ms. The results affirm that Newcomb's T of Eq. (27.1) may be associated with theoretically uniform Ephemeris Time (T_{E}) and interpreted as the independent variable of orbital motion and that the right ascension of Newcomb's fictitious mean sun is recognizable as the "ephemeris mean sun" $R_{\odot}(T_{\text{E}})$.

The Mean Sun in the Twenty-First Century

Any measure of time based on an average of apparent solar time is dependent upon the particular theory adopted for the Sun. Newcomb's convention for the fictitious mean sun was based on a pre-relativistic solar system theory, fitted to telescopic observations from the eighteenth and nineteenth centuries. In principle, a more

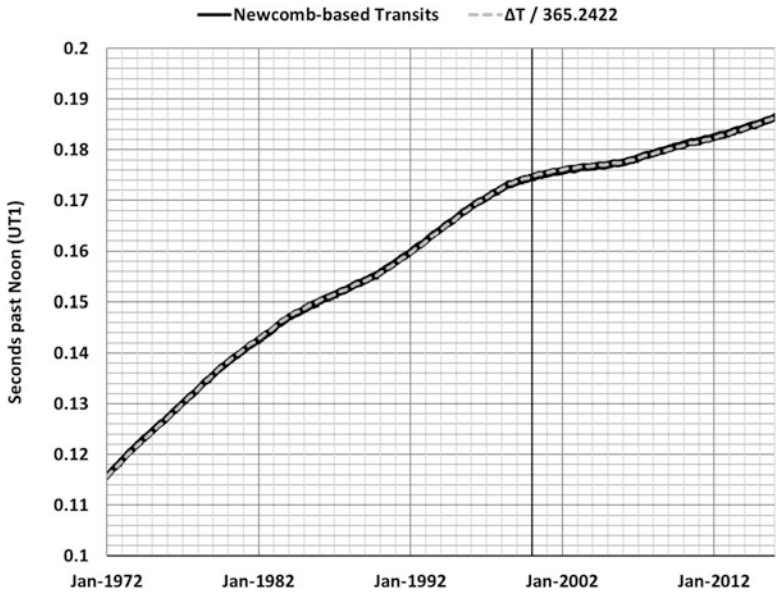


Fig. 27.5 Times of simulated solar transits, adjusted for the equation of time based on Newcomb's expression for right ascension of the fictitious mean sun

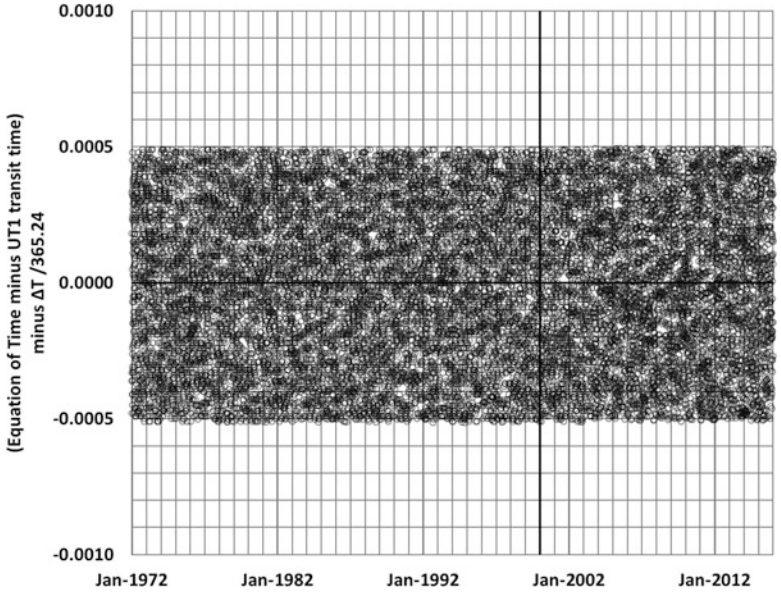


Fig. 27.6 Residuals of Newcomb's T_E-based transits minus $\Delta T/365.2422$. The accuracy of the transit times (UT1) was preserved to only ± 0.5 ms in the simulation

modern convention for mean solar origin and rate should match the long-term behavior of the true Sun more accurately.

The most contemporary expressions for the mean elements of the Earth, of which the authors are aware, are those of the *Variations Séculaires des Orbites Planétaires* (VSOP) 2013 (Simon et al. 2013). These elements were developed from an analytical theory adjusted to the recent European numerical integration IPOP10a and are almost identical to the mean elements of Simon et al. (1994) fitted to JPL Development Ephemerides. VSOP 2013 expresses the mean longitude of the Earth (converted here to arcseconds) as[4]

$$L_{\oplus} = 361679.2260111'' + 1295977422.9089161''T[\text{epoch}\,2000], \qquad (27.24)$$

where T is in units of Julian millennia TDB. To define the origin of right ascension for a fictitious mean sun from the mean longitude of the Earth $L_{\oplus}$, one must add 180 degrees (648,000″) to $L_{\oplus}$ and then adjust for annual aberration, κ, to account for the geocentric frame of reference. A modern value for κ at epoch J2000 is 20.49551″ or 1.36637 s (Astronomical Almanac 2006). Additionally, the dynamical ecliptic and equinox of VSOP 2013 is "inertial" in the sense defined by Standish (1981). An adjustment is therefore necessary to compare results with Newcomb's "rotating" equinox convention, which Standish estimated to be at epoch J2000:

$$\alpha_{\text{rotating}} = \alpha_{\text{inertial}} - 0.09366''. \qquad (27.25)$$

Thus, $361679.2260111'' + 648{,}000'' - 20.49551'' - 0.09366''$ updates $\alpha_0 = L_0 = 1009658.63684''$.

From inspection, the rate term within Eq. (27.24) corresponds to λ_1 in Eq. (27.20), because VSOP 2013 mean elements are with respect to an inertial frame. This requires the addition of precession in longitude, h_1, so that the rate term is compatible with durations of the tropical year. The value of h_1 compatible with IAU 1976 is 50290.966 arcseconds per Julian millennium (Kaplan 1981), so that $L_1 = \lambda_1 + h_1 = 1295977422.909 + 50290.966 = 1296027713.875$. Per Eq. (27.21), a mean sun, whose equinox is affected by general precession in right ascension of IAU 1976 precession, would have its right ascension expressed as

$$\alpha = \alpha_0 + \mu T + \left(46124.362''T + 139.656''T^2 - 0.0927''T^3\right)[\text{epoch}\,2000], \qquad (27.26)$$

where T is in units of Julian millennia TDB from epoch J2000.0. Equating the coefficients of like terms, $\mu = (\lambda_1 + h_1) - m_1 = 1295981589.513$. For this analysis then, an expression for the right ascension of the mean sun, commensurate with the mean elements of VSOP 2013, is

[4] ftp://ftp.imcce.fr/pub/ephem/planets/vsop2013/solution/README.pdf.

$$\alpha_{\text{VSOP 2013}} = 1009658.63684'' + 1296027713.87492''T + 139.656''T^2 - 0.0927''T^3, \quad (27.27)$$

where T is in units of Julian millennia TDB from epoch J2000. This expression has a corresponding annual period of 31556925.181576 s (365.24218960 d) at epoch J2000. The difference between the expressions of right ascension for the mean sun of VSOP 2013 v. Newcomb at J2000 is then

$$\alpha_{\text{VSOP 2013}} = R_{\odot} + 0.41084'' - 8.05508''T + \cdots[\text{epoch 2000}]. \quad (27.28)$$

Thus, to a fair degree of approximation, an expression for mean solar time based on VSOP 2013 mean elements may be represented as

$$\text{GMT}_{\text{VSOP 2013}} \approx \text{UT1} + 0.002738\ \Delta T + (\Delta\alpha_0 + \Delta\mu\ T) \quad [\text{epoch 2000}] \quad (27.29)$$

where $\Delta\alpha_0 \approx 0.02739$ s and $\Delta\mu \approx -0.53701$ s. It is more than a coincidence that $\Delta\alpha_0$ is approximately 0.02 s, which is the degree of separation predicted after 100 years, according to Newcomb's estimate of the secular acceleration differences between his mean sun and the mean longitude of the actual Sun.

The general result based on VSOP 2013 is illustrated in Fig. 27.7, along with Newcomb's result. UTC was included in the figure for scale; the discontinuities in the graph of UTC result from the introduction of *leap seconds* which maintain

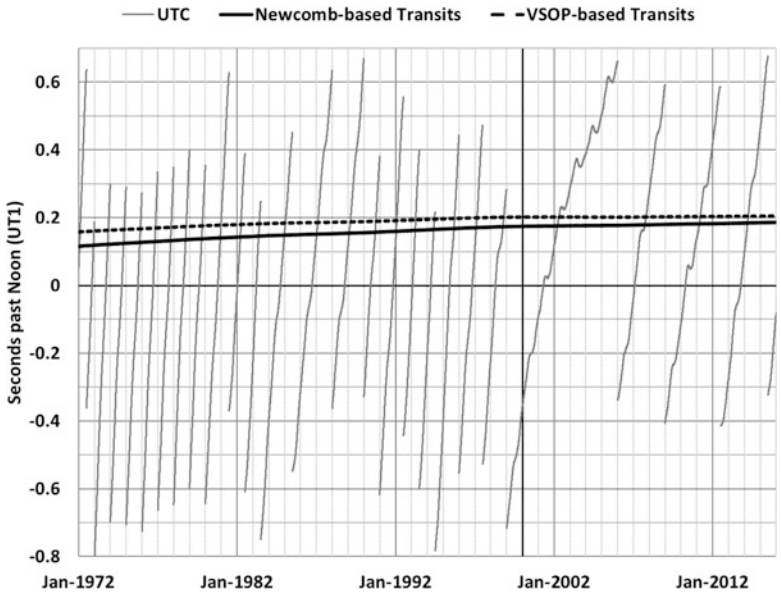


Fig. 27.7 Times of simulated transit after noon UT1 for different conventions of Greenwich mean solar time

UTC's proximity to UT1. The time expressed by either mean sun differs from Universal Time by an amount which is, presently, much less than the variation between UTC and UT1. Of the three realizations illustrated, Newcomb's mean sun is closest to UT1 currently, but the mean sun of VSOP 2013 is expected to be closest after the middle of this century, due to the reduction of rate evidenced in Eq. (27.28).

Conclusion

The difference between UT1, and time indicated by the geometry of Newcomb's mean sun, is a consequence of how Earth rotation has been historically measured. In principle, durations in mean solar time are the combination of sidereal time (Earth rotation) and solar mean motion (the tropical year). Yet for precise work, the measure of Earth rotation (sidereal time) was multiplied by a scale factor representing the ratio of solar time to sidereal time. These two approaches do not yield equivalent results unless Earth's rotation is constant.

In this paper, the right ascension of Newcomb's fictitious mean sun is expressed with respect to the dynamical equinox of epoch J2000, and its behavior is compared with UT1. To make the comparison, software was used to simulate the transit times (UT1) of the apparent Sun at the prime meridian over four decades. The transit times were then compared to the equation of time. The simulation affirmed that the geometry of Newcomb's mean solar time at the prime meridian differs from UT1 by nearly $\Delta T/365.2422$, where ΔT is Terrestrial Time minus UT1. This divergence was anticipated in the mid-twentieth century, when Ephemeris Time was recognized as its own separate scale. The present separation between the geometry of Newcomb's mean solar time and Universal Time is about 0.2 s.

Newcomb's mean sun was intended to move at a uniform rate for the purpose of providing a more theoretically uniform convention for timekeeping, but this expression deviates secularly from the longitude of the true Sun. To better predict how UT1 evolves with respect to the true Sun, expressions based on updated mean elements may be approximated as

$$\mathrm{GMT}_{(\text{analytic mean sun})} - \mathrm{UT1} \approx \Delta T/365.2422 + (\Delta\alpha_0 + \Delta\mu\, T + \cdots) \quad [\text{epoch } 2000]. \tag{27.30}$$

where $\Delta\alpha_0$ and $\Delta\mu$ represent offsets from the origin and rate of the right ascension of Newcomb's mean sun at epoch J2000, T being in Julian millennia TDB, and terms of T^2 are neglected. To compare Newcomb's expression and the UT1 series with an up-to-date solar system theory, the mean solar longitude of VSOP 2013 was considered. VSOP 2013 constants differ from Newcomb's by (0.02739 s–0.53701 s T), after accounting for differences between their conventional frames of reference. The difference in origin is small relative to the currently estimated

divergence of 0.2 s from Universal Time and smaller than the varying duration of the tropical year, which decreases by one-half second per century.

The slight differences between VSOP 2013 and Newcomb's constants affirm that Newcomb's c.1895 values were relatively accurate. The separation of Newcomb's mean sun from UT1 is primarily due to the historical variation of Earth rotation represented by ΔT and secondarily due to the higher-order differences in secular accelerations of the actual Sun (terms of T^2 and higher) compared to an expression for the mean sun in right ascension. These differences in secular accelerations have become non-negligible as T has increased since Newcomb's time. Nevertheless, UTC is allowed to separate from UT1 by 0.9 s, a larger amount than a mean sun based on either Newcomb's constants or the constants of VSOP 2013. Accordingly, UT1 still appears synonymous with the geometry of "mean solar time at the prime meridian" to within a few tens of a second, despite the definition and operational maintenance of Universal Time having evolved significantly over several centuries.

References

Astronomical Almanac for the Year for the Year 2009 (U.S. Gov't Printing Office, Washington, 2006), p. K7

C. Audoin, B. Guinot, *The Measurement of Time—Time, Frequency and the Atomic Clock* (Cambridge University Press, Cambridge, 2001), pp. 265–66

Ball, *Elements of Astronomy*, 1880, p. 377

K.M. Borkowski, The tropical year and the solar calendar. J. R. Astron. Soc. Can. **85**(3), 127 (1991). (Whole No. 630)

N. Capitaine, B. Guinot, D.D. McCarthy, Definition of the celestial ephemeris origin and of UT1 in the International celestial reference frame. Astron. Astrophys. **355**, 398–405 (2000)

W. de Sitter, D. Brouwer, On the system of astronomical constants. Bull. Astron. Inst. Neth. **8**, 219 (1938)

Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac (Her Majesty's Stationery Office, London, 1961), pp. 74–82.

R.M. Green, *Spherical Astronomy* (Cambridge University Press, Cambridge, 1985), pp. 241–242

B. Guinot, Solar time, legal time, time in use. Metrologia **48**, S181–S185 (2011)

Gummere, *An Elementary Treatise on Astronomy in Two Parts*, 3rd edn. (Kimber & Sharpless, Philadelphia, 1842), p. 50

D. Howse, *Greenwich Time and the Longitude* (Philip Wilson Publishers Ltd, London, 1997), pp. 45–49

D.W. Hughes, B.D. Yallop, C.Y. Hohenkerk, The equation of time. Mon. Not. R. Astron. Soc. **238**, 1529–1535 (1989)

G.H. Kaplan, (ed.), *The IAU Resolutions on Astronomical Constants, Time Scales, and the Fundamental Reference Frame*. USNO Circular No. 163 (U.S. Naval Observatory, Washington DC, 1981), December 10, 1981, p. B5

S. Malys, J.H. Seago, N. Pavlis, P.K. Seidelmann, G.H. Kaplan, Why the Greenwich Meridian moved. J. Geod. **89**(12), 1263–1272 (2015)

N. Maskelyne (1764), The Monthly Review; or Literary Journal: by Several Hands. vol 34, pp. 228–229

J. Meeus, D. Savoie, The history of the tropical year. J. Br. Astron. Assoc. **102**(1), 40–42 (1992)

O. Neugebauer, *A History of Ancient Mathematical Astronomy*, vol 1 (Springer, Berlin, 1975), pp. 61–66

S. Newcomb, E.S. Holden, *Astronomy for Students and General Readers*, Revised 2nd edn. (1880), p. 260

S. Newcomb, On the present state of the theories of the celestial motions. Sidereal Messenger **2**(2), 39 (1883)

S. Newcomb, *The Elements of the Four Inner Planets and the Fundamental Constants of Astronomy*, Supplement to The American Ephemeris for 1897, 1895a, p. 188

S. Newcomb, Inequalities of long period: and of the second order as to the masses, in the mean longitudes of the four inner planets. in *From astronomical papers prepared for the use of the American ephemeris and Nautical Almanac*, vol 5, Part II (U.S. Nautical Almanac Office, Washington, 1895b)

S. Newcomb, Tables of the motion of the earth on its axis and around the sun, in *From Tables of the Four Inner Planets. Astronomical papers prepared for the use of the American ephemeris and Nautical Almanac*, 2nd edn. vol VI, Part I, 1898, p. 9

S. Newcomb, *A Compendium of Spherical Astronomy* (Macmillan Company, 1906), p. 120

H.C.F.C. Schjellerup, Recherches sur l'Astronomie des Anciens—II. On the total eclipses observed in China in the years B.C. 708, B.C. 600, and B.C. 548. Copernicus **I**, 41–47 (1881)

J.H. Seago, P.K. Seidelmann, The mean-solar-time origin of universal time and UTC. Paper AAS 13–486 from Tanygin et al. (ed., 2013), *Space Flight Mechanics 2013 – Advances in the Astronautical Sciences*, Proceedings of the 23rd AAS/AIAA Space Flight Mechanics Meeting, (Kauai, Hawaii, 2013), vol 148, Part 2, pp. 1789–1807

J.H. Seago, R.L. Seaman, P.K. Seidelmann, S.L. Allen (ed.), *Requirements for UTC and civil timekeeping on earth*. A Colloquium Addressing a Continuous Time Standard. AAS Science and Technology Series, vol 115 (Univelt, Inc., San Diego, 2013)

J.H. Seago, R.L. Seaman, S.L. Allen ed. *Decoupling civil timekeeping and earth rotation*. A Colloquium Exploring Implications of Redefining Coordinated Universal Time (UTC). AAS Science and Technology Series, vol 113 (Univelt, Inc., San Diego, 2011)

P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac* (University Science Books, Mill Valley, 1992), p. 78

J.-L. Simon, G. Francou, A. Fienga, H. Manche, New analytical planetary theories VSOP2013 and TOP2013. Astron. Astrophys. **557**, A49 (2013)

J.-L. Simon, P. Bretagnon, J. Chapront, M. Chapront-Touze, G. Francou, J. Laskar, Numerical expressions for precession formulae and mean elements for the Moon and the planets. Astron. Astrophys. **282**(2), 663–683 (1994)

N. Sivin, *Granting the Seasons: The Chinese Astronomical Reform of 1280, with a Study of Its Many Dimensions and an Annotated Translation of Its Record* (Springer, New York, 2009), p. 206

C.A. Smith, G.H. Kaplan, J.A. Hughes, P.K. Seidelmann, B.D. Yallop, C.Y. Hohenkerk, Mean and apparent place computations in the new IAU system. I. The transformation of astrometric catalog systems to the equinox J2000.0. Astron. J. **97**(1), 266 (1989)

E.M. Standish, Two differing definitions of the dynamical equinox and the mean obliquity. Astron. Astrophys. **101**, L17–L18 (1981)

D. Steel, *Marking time – The epic quest to invent the perfect calendar* (Wiley, New York, 2000), pp. 381–384

E.J. Stone et al., 1883 April 13 meeting of the Royal Astronomical Society. Observatory **6**, 140 (1883)

S. E. Urban, P. K. Seidelmann (eds.), *Explanatory Supplement to the Astronomical Almanac*, 3rd edn. (University Science Books, Mill Valley, 2012), p. 81

R. Woodhouse, *An Elementary Treatise on Astronomy* (Cambridge University Press, Cambridge, 1812)

E.W. Woolard, G.M. Clemence, *Spherical astronomy* (Academic Press, New York, 1966), pp. 346–347

Chapter 28
How Gravity and Continuity in UT1 Moved the Greenwich Meridian

Stephen Malys, John H. Seago, Nikolaos K. Pavlis, P. Kenneth Seidelmann, and George H. Kaplan

Abstract The concept of "Greenwich Mean Time," and its modern equivalent, Universal Time, is based on the changing angle between a "prime meridian" and some point on the celestial sphere. In 1884, at the International Meridian Conference, it was recommended that the prime meridian "to be employed as a common zero of longitude and standard of time-reckoning throughout the globe" pass through the Airy "transit instrument at the Observatory of Greenwich." Today, observatory visitors must walk approximately 102 m east before their satellite-navigation receivers indicate zero longitude. The need to maintain continuity in Universal Time by the Bureau International de l'Heure, in conjunction with a transition from astronomical to geodetic coordinates by 1984, when optical astronomical methods were replaced by modern techniques, is responsible for this offset. The difference between astronomical and geodetic longitudes is the deflection of the vertical in the east-west direction. Modern techniques enabled the establishment of a global reference frame centered at the center of mass of the Earth, through which the plane of the geodetic prime meridian passes. While the geodetic prime meridian does not intersect the Airy transit instrument at the surface of the Earth, its orientation with respect to the celestial sphere has remained intact.

Keywords Greenwich meridian • Deflection of the vertical • Prime meridian • UT1 • ITRF zero meridian • IERS reference meridian • Greenwich mean time

S. Malys (✉) • N.K. Pavlis
National Geospatial-Intelligence Agency, 7500 GEOINT Dr., Springfield, VA 22150, USA
e-mail: stephen.malys@nga.mil

J.H. Seago
Analytical Graphics, Inc., 220 Valley Creek Blvd, Exton, PA 19341-2380, USA

P.K. Seidelmann
Astronomy Department, University of Virginia, P.O. Box 400325, Charlottesville, VA 22904, USA

G.H. Kaplan
United States Naval Observatory (contractor), 3450 Massachusetts Avenue, NW, Washington, DC 20392, USA

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_28

The Situation at Greenwich

The Airy Transit Circle is a nineteenth-century telescopic instrument at the Royal Observatory, Greenwich, famous as the origin for global longitude. The observatory is now a museum with signage marking a line in the pavement as the "Prime Meridian of the World." However, its longitude is 00° 00′ 05.3″W (Howse 1997) in the International Terrestrial Reference Frame (ITRF), and in the World Geodetic System (WGS 84) used by the US Global Positioning System (GPS), a fact that is widely noticed by tourists with satellite-navigation receivers. The WGS 84 and the ITRF zero-longitude meridian plane is located approximately 102 m east of the telescope, as shown in Fig. 28.1. The apparent discrepancy has raised questions of how and when it arose and whether the worldwide system of longitude systematically shifted by this amount (Malys et al. 2015).

Fundamental optical astrometry makes use of the direction of gravity at individual telescope sites to define the astronomical latitude and longitude of observatories. Astronomical longitude is the angle between two conventional astronomical-meridian planes. The plane of the local astronomical meridian contains the local meridian defined by the great circle passing through the celestial poles and the local direction of astronomical zenith (along the local plumb line). The conventional plane of reference that was adopted in 1884 is the astronomical-meridian plane of the Airy transit instrument at the Observatory at Greenwich.

Geodetic longitude, on the other hand, makes use of a geodetic meridian plane that contains the minor axis of an adopted reference ellipsoid and the direction of the normal to that reference ellipsoid at the site. Reference ellipsoids are mathematical constructions designed to fit the surface of the Earth either globally or locally. The angle between the plumb line and the normal to the ellipsoid at a site is called the deflection of the vertical (DoV), described in further detail below, in section "Regional Versus Global Geodesy".

Optical Methods of Determining Longitude and Earth Rotation

Prior to the 1930s, the most accurate timekeeping was astronomical, determined by optical measurements of Earth rotation relative to stars. The traditional *transit circle* or *transit instrument* is a special-purpose refracting telescope, mounted so that the optical axis is confined to rotate within the plane of the local *astronomical* meridian. It was developed by Roemer near the end of the seventeenth century (Grant 1852), and the technology was well established when the seventh Astronomer Royal, Sir George Airy, installed his namesake instrument at Greenwich in 1850, 0.3″ (6 m) east of Bradley's meridian (shown in Fig. 28.1)—an earlier meridian still used within the UK by Great Britain's Ordnance Survey for official mapping purposes (MacDonald 1985).

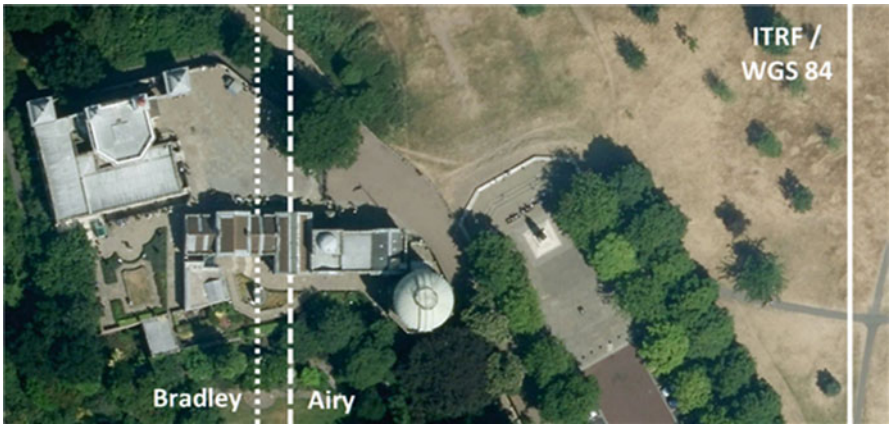


Fig. 28.1 The Bradley and Airy meridians (*dotted lines*) and the ITRF Zero Meridian (*solid line*). Imagery © 2014 Infoterra Ltd & Bluesky

A transit instrument is oriented relative to the local gravity field, as depicted in Fig. 28.2. During the nineteenth century, a transit instrument's axis of rotation would have been corrected according to the readings of a striding spirit level (**L** in the figure), with the pivots (**V**) vertically adjusted to remove detectable inclination as necessary. Local gravity was further used to align the collimation of the optical train with astronomical zenith, by placing a basin of mercury beneath the telescope (**Hg** in the figure). The offset between an illuminated eyepiece reticle (**E**) and its reflection in the mercury basin was recorded nightly for later correction (Christie 1903). The measured meridian plane did not necessarily pass through the center of mass of the Earth, because the local plumb line could be tilted in the east-west direction with respect to a line connecting the site and the center of mass of the Earth (see also Fig. 28.6).

Transit instruments were used to precisely observe the passage of specially cataloged "clock stars" of known right ascension; local sidereal time was indicated when these stars coincided with the calibrated reticle within the telescope. After the observed sidereal time was converted to mean solar time, according to a conventional relationship, observatory clocks and time signals were adjusted to match the astronomical observations. The difference between astronomically determined local mean times t (in hours) is proportional to the difference in local longitudes Λ (in degrees):

$$t - t_0 = (\Lambda - \Lambda_0)/15, \tag{28.1}$$

where Λ_0 and t_0 refer to the astronomical longitude and local mean time of the reference (usually prime) meridian.

At the International Meridian Conference of 1884, it was recommended that the prime meridian pass through the Greenwich transit instrument (Nautical Almanac Office 1961), as the Greenwich meridian was already a de facto standard for

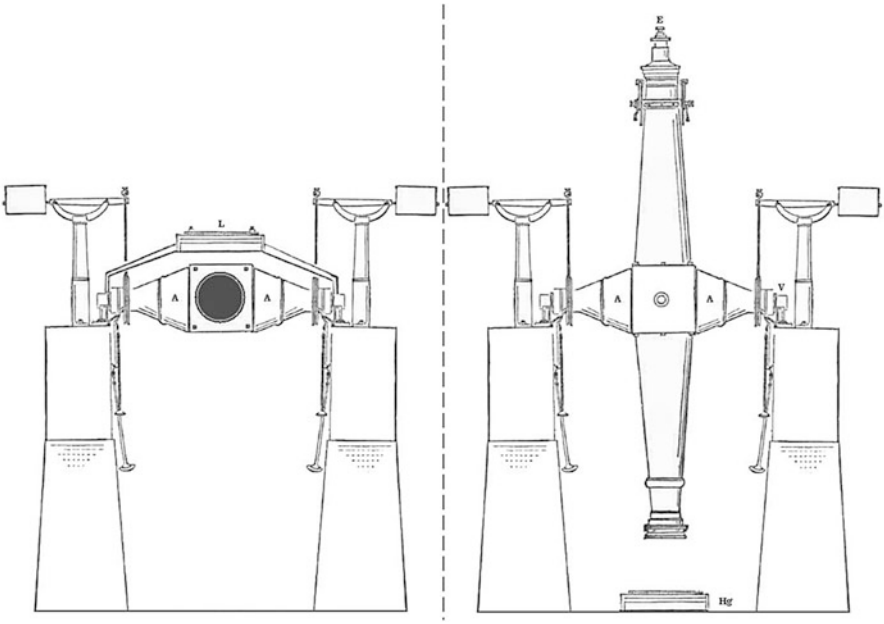


Fig. 28.2 Leveling and collimation of a nineteenth-century transit instrument. Transit in a horizontal position (*left*) and looking toward nadir (*right*) (Norton 1872)

navigation before then. Therefore, the mean solar time determined by the Airy Transit Circle—Greenwich Mean Time (GMT)—became an internationally recognized basis for global timekeeping and navigation. As per Eq. (28.1), any astronomical time service could therefore convert its own determination of local mean time (t) to GMT (t_0) by subtracting the longitude difference from Greenwich ($\Lambda - \Lambda_0$). But after a comparison of worldwide radio broadcasts revealed that time signals differed by up to several seconds, the *Bureau International de l'Heure* (BIH) was established at the Paris Observatory in 1912 to help coordinate the national time bureaus (Guinot 2000).

From 1920 to 1928, the BIH maintained its own version of GMT, designated *heure demi-definitive*, by retarding its local determinations of Paris mean time by approximately $9^{m}21^{s}$ to account for the longitude difference between Paris and Greenwich. Meanwhile at Greenwich, the east pier of the Airy instrument began to noticeably sink relative to the west pier after a very dry summer in 1921. More frequent leveling adjustments were required, with new pivot bearings installed in 1924. When a small reversible transit instrument was brought into supplemental use at Greenwich in 1926, it gave noticeably different results from the Airy Transit Circle. A comparison of the observations with other time signals suggested that the error resided mostly in the collimation of Airy's transit, such that the British Time Service transitioned its timekeeping program to the small transit after July 1927 (Gething 1954). However, the supplementary instruments at Greenwich affiliated

with the Royal Greenwich Observatory (RGO) continued to be assigned a conventional longitude of exactly zero by the BIH (BIH 1933 and subsequent), their observations being adjusted for their longitude separation from the Airy Transit.

During this period, the newly formed International Astronomical Union (IAU) dissuaded astronomical usage of the term "Greenwich Mean Time" and its abbreviation "GMT" because of a change in almanac conventions starting in 1925 and began recommending the term "Universal Time" (UT) instead. Starting in 1929, the BIH introduced *heure definitive* based on the contributed time observations of six observatories, a number which gradually increased to 37 by 1963. These contributing institutions maintained their own versions of UT and the BIH estimated definitive Universal Time, using a weighted average of time-service signals. For purposes of global timekeeping by the BIH then, the Greenwich Observatory was superseded by the concept of the "mean observatory," to which Greenwich became one of many contributors. The formation of this mean observatory involved adopted corrections to the contributed data and weights that could change annually (Feissel 1980). These corrections could account for perceived errors in the signals of individual time bureaus, including changes in the adopted positions of the observed stars and possible changes in the adopted astronomical longitudes of the observing sites. To maintain a continuous and consistent series of UT, the BIH also assumed the responsibility for maintaining a terrestrial reference system used for the analysis (Guinot 2000). The mean longitudes and latitudes of contributing stations were estimated from their astronomical observations, such that the early BIH terrestrial reference frames were strictly with respect to astronomical coordinates (Mueller 1969).

Improved timekeeping technology provided the opportunity in the 1930s to use the astronomical observations for measuring and understanding irregularities in the length of day (Moritz and Mueller 1987). This resulted in the discovery of a seasonal variation in UT, which was corrected in UT broadcasts, and the corrected time scale was denoted later as UT2. Coincident with the development of atomic clocks in the 1950s, the astronomical community began to distinguish between time scales based on measures other than Earth rotation (e.g., dynamical timescales, such as Ephemeris Time).

By the 1950s, traditional transit and meridian circles were being replaced by zenith telescopes, which pointed directly overhead, and smaller circumzenithal instruments that surveyed all directions at a fixed angle from the zenith. Circumzenithal design was perfected with the development of the Danjon impersonal astrolabe. During the International Geophysical Year (1957–1958), ten photographic zenith tubes and 16 Danjon astrolabes were known to be in operation (Munk and MacDonald 1975). These newer optical designs also relied on an artificial horizon (a mercury surface) to reflect the starlight directly and were therefore self-leveling with respect to local gravity. The estimated probable error of either a photographic zenith tube or Danjon astrolabe over one night was about ± 4 ms ($0.06''$), compared to ± 18 ms for a small transit employing an impersonal micrometer (Nautical Almanac Office 1961, p. 448).

The BIH also corrected the astronomically determined time for small, periodic variations in the location of the rotational pole, monitored by the International Latitude Service (ILS) since 1899 (Moritz and Mueller 1987, p. 397). As per IAU recommendations, UT corrected for polar motion was designated by time services as UT1, whereas uncorrected determinations were denoted as UT0. After the BIH adopted various reference poles of different epochs during the 1950s and 1960s (Robbins 1967), the International Union of Geodesy and Geophysics (IUGG) and IAU (IUGG 1967, IAU 1968) resolved that the BIH refer to the average ILS pole from 1900 to 1905—known as the Conventional International Origin—as part of a broader update to celestial and terrestrial reference systems. To preserve the continuity of UT1–UT0 for each contributing station within this new "1968 BIH System," the adopted longitude, Λ_i, of each observatory was systematically changed (BIH 1969) according to

$$\Delta\Lambda_i = -0.007'' \sin(\Lambda_i) \tan(\Phi_i) + 0.233'' \cos(\Lambda_i) \tan(\Phi_i) \tag{28.2}$$

where Φ_i is the latitude. Although it was no longer operational by that time, the longitude of the Airy Transit Circle would have changed from zero to 0.2927″ W (about 6 m) within this new system (Schmid 1974; MacDonald 1985).

In 1973, the BIH began incorporating estimates of polar motion from the US Navy TRANSIT (satellite Doppler) navigation system into its own processes (Guinot 1979). Other experimental techniques, such as connected-element radio interferometry, very-long-baseline interferometry (VLBI), and lunar laser ranging (LLR), were gradually introduced as a means to estimate UT1, along with the traditional optical measurements. A joint IAU- and IUGG-sponsored program, called MERIT-COTES, evaluated the accuracy of these new sources (Wilkins 2000), affirming their advantages for Earth orientation and geodesy over traditional astrometry. Beginning in 1978, these new data were averaged into BIH determinations of UT1 (Fig. 28.3).

By 1984, the BIH discounted optical astrometry completely and established a new terrestrial coordinate system designated as the "BIH Terrestrial System," or BTS 84 (BIH 1985). This system was based on weighted solutions of Earth-orientation parameters from analysis centers and networks of VLBI, SLR, and LLR geodetic stations that were independent of the former network of optical systems (Boucher and Altamimi 1985, 1986). The orientation of the network was constrained to match the ensemble of Earth-orientation measurements made from 1980 through 1983. Thus, the series of UT1 determinations based on modern, nontraditional techniques maintained continuity with the former series dominated by optical astrometry, within the uncertainty of the optical measurements. This also aligned the zero meridian of the geodetic station network with the direction of the astronomical zero meridian of the former series.

By 1988, the BIH and the International Polar Motion Service (successor to the ILS) had been replaced by a single entity, the International Earth Rotation Service (IERS 2014; Wilkins 2000), which assumed responsibility for maintaining the terrestrial and celestial reference frames consistent with the Earth-orientation

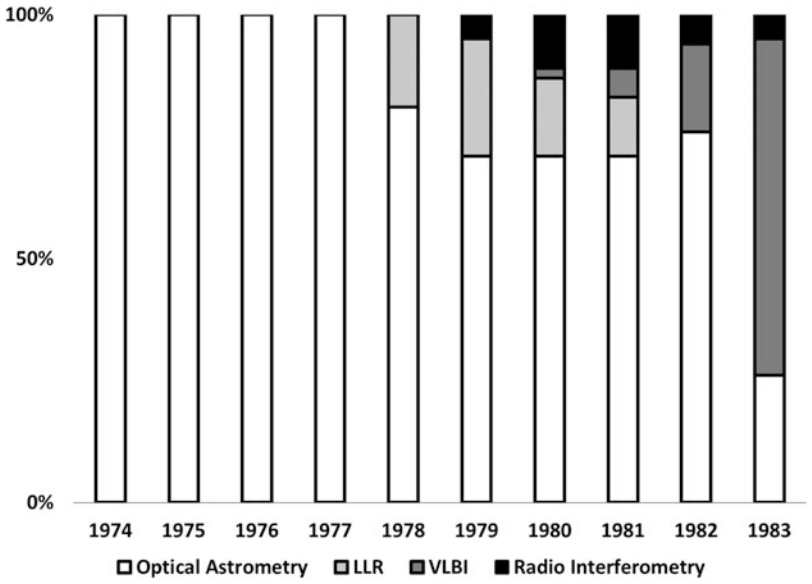


Fig. 28.3 Relative weights for BIH combined solutions for UT1, 1974–1983 (Yatskiv 1985)

parameters it published routinely. The subsequent realizations of the International Terrestrial Reference Frame (ITRF) by the IERS preserved the orientations relative to the reference meridian of the BTS (see Jekeli 2006, pp. 3–14). Modern geodetic reference frames, such as the ITRF, are defined by the Cartesian position and velocity coordinates of an ensemble of stations in three dimensions; the corresponding reference (zero) meridian is the Earth-fixed X-Z plane. The IERS Conventions (IERS 2010) recommend the term "ITRF zero meridian" for this plane, although the terms *IERS Reference Meridian* and *International Reference Meridian* (IRM) are also in use. A timeline of notable events in the history of Universal Time and the Greenwich meridian is shown in Fig. 28.4.

Regional Versus Global Geodesy

Any point on or above the surface of the Earth can be defined by a triad of coordinates: the astronomical latitude Φ, astronomical longitude Λ, and the height above or below a surface of equal potential of the Earth's gravity field. The angle between the direction of the gravity vector (the vertical direction realized by a builder's plumb line) and a line passing through the point parallel to the rotation axis of the Earth defines the complement of the astronomical latitude, $90^\circ - \Phi$. The plane defined by these two lines is the astronomical-meridian plane passing through the point. In general, the geocenter does *not* lie in the astronomical-meridian plane of an arbitrary point. The angle between this meridian plane and the prime meridian

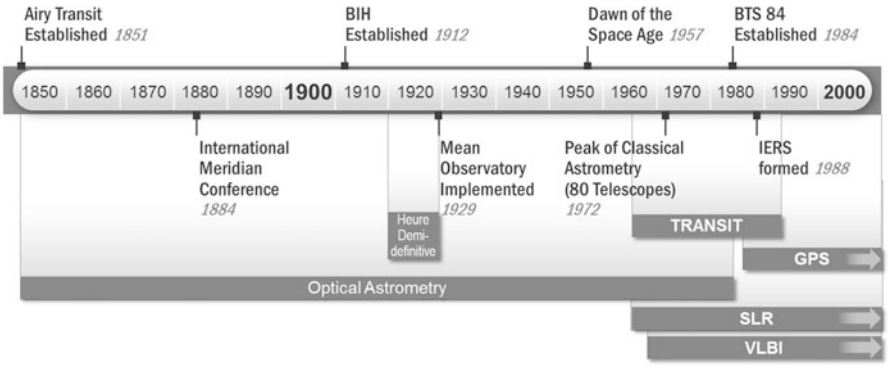


Fig. 28.4 Events in the history of Universal Time and the Greenwich meridian

defines the astronomical longitude, Λ, of the point. A suitable reference surface for the definition of the orthometric height, H, is an equipotential surface of the Earth's gravity field—the geoid. The geoid is closely related to the world's ocean surface, after the latter is adjusted for time-varying effects (e.g., ocean tides) and averaged over long periods. The coordinates Φ, Λ, and H are called *natural* coordinates (Heiskanen and Moritz 1967, section 2–4), since their definition is possible using naturally occurring surfaces and directions. These surfaces and directions are inextricably linked to the Earth's gravity field and to the Earth's rotation.

Advances in the high-resolution determination of the Earth's global gravity field have enabled the estimation of the best-fitting ellipsoids of revolution approximating the global figure of the Earth. These mean-Earth ellipsoids are centered on the center of mass of the Earth (the geocenter), and their axes are aligned with the axes of three-dimensional, global coordinate systems. In these geodetic coordinate systems, the plane of the longitude meridian passing through a given location includes the minor axis of the ellipsoid of revolution (north pole to south pole) and is orthogonal to the surface of the ellipsoid. Height is measured from the ellipsoid. The projection of *geodetic* coordinates onto an ellipsoid of revolution is more suitable for cartographic applications than natural coordinates, because an ellipsoidal surface can be projected onto a flat surface with (sometimes complicated) mathematical formulas. Defining a *datum*, composed of a reference ellipsoid whose size, shape, position, and orientation are carefully chosen to support efficient mapping over a given region, was a primary endeavor of traditional geodesy. For a *local* or *regional datum*, it was essential that the surface of the chosen reference ellipsoid conform to the local or regional geoid as well as possible. The North American Datum 1927 (NAD27) and the European Datum 1950 (ED50) are widely known examples of regional datums. Such regional mapping applications do not require the origin of the coordinate system to be at the geocenter. Moreover, a practical technique for precisely locating the geocenter relative to the Earth's surface was not available until space-geodetic positioning techniques were enabled

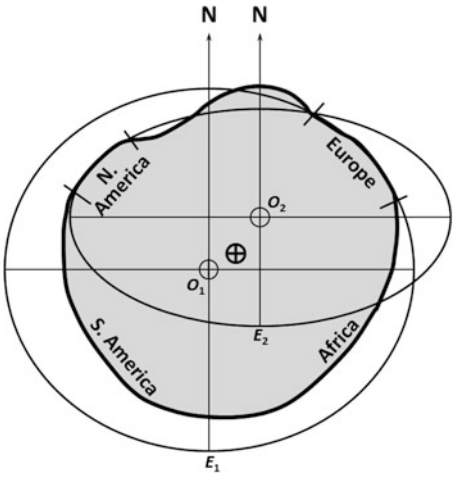


Fig. 28.5 Spatial relationship of the geoid with two "regionally best-fitting" ellipsoids (not to scale). Ellipsoid 1 (E_1) fits Europe, while Ellipsoid 2 (E_2) fits North America. Neither ellipsoid is geocentric nor provides an ideal global fit of the figure of the Earth (Fischer 1975)

through the use of artificial satellites. Figure 28.5 schematically illustrates this situation.

The natural coordinates [Φ, Λ, H] are related to geodetic latitude, φ; longitude, λ; and ellipsoidal height, h, by

$$\left.\begin{array}{ll} \xi = \Phi - \varphi & \text{(a)} \\ \eta = (\Lambda - \lambda)\cos\varphi & \text{(b)} \\ N = h - H & \text{(c)} \end{array}\right\} \tag{28.3}$$

where ξ is the meridional component of the DoV, η is the prime vertical component of the DoV, and N is the geoid undulation, i.e., the height of the geoid above (positive N) or below (negative N) the surface of the reference ellipsoid (Heiskanen and Moritz 1967). These equations relate the geodetic quantities [φ, λ, h] to the physical quantities [Φ, Λ, H] associated with the gravity field; Eq. (28.3b) assumes a common origin of longitude. The DoV components are shown graphically in Fig. 28.6.

If the geodetic coordinates [φ, λ, h] of a point have been determined with respect to a given datum, and if the natural coordinates [Φ, Λ, H] have been determined also (e.g., from astronomical observations and from spirit leveling), then one can obtain the deflection components (ξ_{astro}, η_{astro}) and the geoid undulation (N) at that point from Eqs. (28.3). Such a *geometric* determination can be realized by making astronomical observations of Φ and Λ at a point whose geodetic coordinates [φ, λ, h] have been determined from satellite positioning, and whose orthometric height, H, has been determined from spirit leveling. However, the *astrogeodetic* deflection components (ξ_{astro}, η_{astro}) are specific to the particular datum to which the geodetic coordinates φ and λ refer.

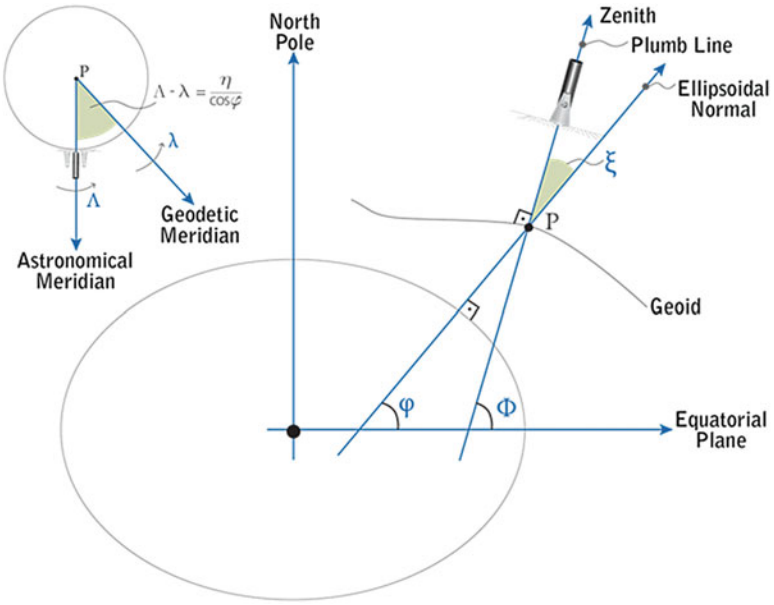


Fig. 28.6 Astronomic (Φ, Λ) vs. geodetic (φ, λ) coordinates, deflection of the vertical in the meridian (ξ) and in the prime vertical (η)

The DoV components (ξ, η) also represent the slopes of the geoid surface with respect to the surface of an *equipotential* reference ellipsoid along the meridian and the prime vertical, respectively (Heiskanen and Moritz 1967, section 2–22):

$$\left.\begin{aligned} \xi_{\text{grav}} &= -\frac{\partial N}{R\partial\varphi} \qquad (a) \\ \eta_{\text{grav}} &= -\frac{\partial N}{R\cos\varphi\partial\lambda} \qquad (b) \end{aligned}\right\} \tag{28.4}$$

where R is a mean-Earth radius. The sign convention in Eqs. (28.3) and (28.4) is such that ξ is positive when the astronomical zenith, Z_a, is north of the geodetic zenith, Z_g, and η is positive when Z_a is east of Z_g (Torge 2001, p. 219). These slopes can be determined *gravimetrically*, either using the integral formulas of Vening Meinesz (Heiskanen and Moritz 1967, section 2–22) or using a spherical harmonic representation of the Earth's gravitational potential. Similarly, the geoid undulation, N, can be obtained gravimetrically, either using Stokes' integral formula (Heiskanen and Moritz 1967, section 2–16) or using again a spherical harmonic representation of the Earth's gravitational potential.

Astrogeodetic and Gravimetric DoV Comparisons

The deflections of the vertical are heavily influenced by the topography surrounding the computation point and by the local irregularities of the gravity field. The Earth Gravitational Model 2008 (EGM2008) is a gravitational model which is complete to spherical harmonic degree 2190 and order 2159 (Pavlis et al. 2012). Such degree and order carry enough resolution to support the determination of gravimetric deflections with adequate accuracy where the terrain is not too rough (so that its omission error is small) and where dense, well-distributed, and accurate terrestrial gravity data are available for the development of the model (so that its commission error is also small). The spatial resolution of EGM2008 is demonstrated to be 10 km; Greenwich is an area with little topographic influence; therefore, this area has little high-spatial frequency content in the local gravity field. It is therefore meaningful to use the gravimetric deflections derived from EGM2008 to estimate the astronomical coordinates for the Airy Transit Circle from a set of well-defined geodetic coordinates of the same location by

$$\left.\begin{aligned} \Phi_{\text{grav}} &= \varphi + \xi_{\text{grav}} && (a) \\ \Lambda_{\text{grav}} &= \lambda + \eta_{\text{grav}} \cdot \sec\varphi && (b) \end{aligned}\right\} \tag{28.5}$$

These "synthetic" astronomical coordinates can then be compared to independent astronomical observations. If the orientation of the zero meridian plane is the same for both astronomical and geodetic longitudes, Eq. (28.3b) holds true (Bomford 1980, p. 100).

Malys et al. (2015) used the EGM2008 spherical harmonic coefficients to degree 2190 and calculated the deflections at Greenwich via harmonic synthesis to obtain $\xi_{\text{grav}} = 2.156''$ and $\eta_{\text{grav}} \cdot \sec\varphi = 5.502''$, a result which agrees with the estimates of $2.15''$ and $5.51''$ made by Ekman and Ågren (2010) via numerical differentiation of the gridded version of the EGM2008 geoid. The results predict an astronomical longitude of $\Lambda_{\text{grav}} = 00^\circ\ 00'\ 00.19''$ E for the Airy Transit Circle; this value is in good agreement with the originally adopted value of zero, it being within the estimated EGM2008 one-sigma commission error, $\pm 0.295''$ (see also Pavlis et al. 2012, section 5).

The astrogeodetic deflection η_{astro} from Eq. (28.3b) should agree with its corresponding gravimetric value η_{grav} from a truncated spherical harmonic series representation corresponding to Eq. (28.4), if the astronomical and geodetic origins of longitude are parallel planes. This condition is not guaranteed in principle (see also Tscherning 1986), but it can be tested where gravimetric determinations of η have been made and where astronomical and geodetic longitudes Λ and λ have been measured. The difference $\Delta\eta = \eta_{\text{astro}} - \eta_{\text{grav}}$ includes errors in the determination of the coordinates, as well as errors of commission and omission in the gravimetric deflection; for a meaningful test, these errors should be small relative to the magnitude of any supposed misalignment of the longitude origin. For an average value of $\Delta\eta$ based on a large sample distributed over a wide area of the Earth,

random (or pseudorandom) errors contributing to each $\Delta\eta$, such as omission errors in gravimetric deflection, will tend to cancel out. The average $\Delta\eta$ will therefore tend toward zero, whenever the longitude origins of η_{astro} and η_{grav} are the same.

Jekeli (1999) reported the results of a comparison at 3561 points distributed over the contiguous United States (CONUS), employing the EGM96 (Lemoine et al. 1998) gravitational model to degree and order 360 to estimate the gravimetric deflections. Pavlis et al. (2008, 2012) reported results employing EGM2008 to degree 2190 and order 2159 over the same 3561 CONUS points, plus 1080 points distributed over Australia. Hirt et al. (2010) reported results based on 1056 points distributed (nonuniformly) over Europe employing EGM2008 for the gravimetric deflections, either alone or augmented by the vertical deflections implied by a residual terrain model (RTM, Forsberg 1984) to reduce omission errors. Table 28.1 presents the mean differences between astrogeodetic and gravimetric deflection, $\Delta\eta$, reported in these studies. The mean differences are within their expected errors (see also Pavlis et al. 2012; Hirt et al. 2010) indicating no regional misalignments between the astronomical and geodetic origins of longitude between the EGM2008 model and the ITRF.

A slightly different approach follows from Ekman and Ågren (2010), who "synthesized" astronomical coordinates as per Eq. (28.5) and then compared values of Λ_{grav} to the historical astronomical longitudes (Λ). For fundamental observatories at Stockholm, København, and Greenwich, their results indicated differences between Λ and Λ_{grav} of $0.4''$, $0.9''$, and $0.2''$, respectively. Malys et al. (2015) extended this approach by considering a comprehensive set of BIH observatory sites. The results provide evidence that the geodetic and astronomical coordinates, and the EGM2008 gravitational model, all share a similar directional origin for longitude.

Summary and Conclusions

When an optical instrument is "leveled," even to the highest accuracy achievable, its vertical orientation is controlled by the direction of the gravity field at its location. The plane of the Airy Transit Circle realized the astronomical prime meridian at Greenwich, precisely oriented to the local direction of gravity. Because of the deflection of the vertical, this meridian did not pass through the geocenter. When optical systems were finally retired from Earth-orientation service by the BIH in 1984, the BIH continued its measurement series for the Earth's rotation by modern geodetic techniques (SLR, VLBI, and LLR) but required continuity in the determinations. These measurements included those for astronomical time, UT1. Requiring continuity in UT1 was equivalent to requiring that the plane of the prime meridian keep its orientation, relative to the celestial sphere, as a function of time (Fig. 28.7). Because the plane of the resulting geodetic coordinates also passes through the center of mass of the Earth, the new prime meridian cuts the Earth's

Table 28.1 Average value of $\Delta\eta = \eta_{astro} - \eta_{grav}$ reported in three studies

Study	Region	No. points	Model	N_{max}	Mean $\Delta\eta$ ($''$)
Jekeli (1999)	CONUS	3561	EGM96	360	−0.02
Pavlis et al. (2008, 2012)	CONUS	3561	EGM2008	2190	−0.16
Pavlis et al. (2008, 2012)	Australia	1080	EGM2008	2190	−0.10
Hirt et al. (2010)	Europe	1056	EGM2008	2190	0.33
Hirt et al. (2010)	Europe	1056	EGM2008+RTM	2190	0.15

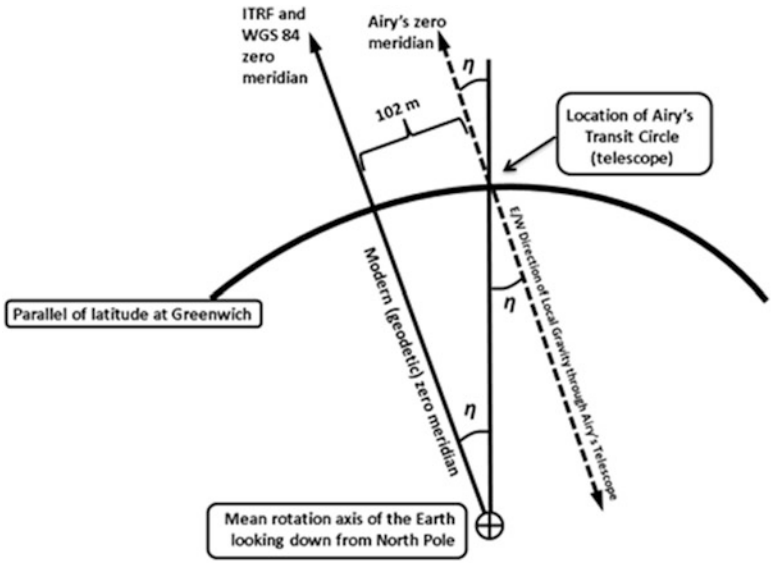


Fig. 28.7 Orientation used to measure Universal Time (UT) remains the same. The dashed meridian line and the solid meridian line are parallel

surface 102 m east of the Airy Transit Circle, an offset which is correctly predicted by EGM2008.

The geodetic reference system established by the BIH in 1984, based on the new techniques, BTS 84, is the ancestor of all the global geodetic systems now in use. The orientation of WGS 84 and the successive realizations of the ITRF were all constrained to maintain continuity in orientation with BTS 84. Because GPS is based on WGS 84, coordinates given by GPS receivers at Greenwich are based on the geodetic prime meridian. As a result, visitors to Greenwich must walk 102 m east of the Airy Transit Circle to find the geodetic longitude zero. This particular distance offset is due to the local DoV value at Greenwich—other observatory sites have different lateral offsets (which may also differ in sign) between the geodetic and astronomical longitude systems. Yet, there is no systematic rotation of global longitudes between the former astronomical system and the current geodetic system. The lack of such a rotation was confirmed by an analysis of observatory coordinates around the world (Malys et al. 2015).

Due to satellite-tracking techniques, the location of the center of mass of the Earth is now known with an accuracy of about 1 cm in three dimensions. In 1984, its location was known with an accuracy of about 1 m. When Sir George Airy established his transit circle in 1850, knowledge of the size, shape, and center of mass of the Earth was limited to hundreds of meters. The trace of its historic meridian line on the surface of the Earth and the location of zero longitude indicated by GPS receivers are both consistent representations within their own conventions, and the offset between them does not indicate an error with the designation of either location.

References

BIH (1933) Bureau International de l'Heure Bulletin Horaire 10 Aug 1933, p. 38

BIH (1969) Bureau International de l'Heure Annual Report for 1968, Observatoire de Paris

BIH (1985) Bureau International de l'Heure Annual Report for 1984, Observatoire de Paris

G. Bomford, *Geodesy*, 4th edn. (Clarendon Press, Oxford, 1980)

C. Boucher, Z. Altamimi, in *Towards an improved realization of the BIH terrestrial frame*, ed. by I.I. Mueller. Proceedings of the International Conference on Earth Rotation and Reference Frames, 31 July–2 August,1985, Columbus, Ohio, USA, MERIT/COTES Report, vol. 2 (Ohio State University, Columbus, 1985)

C. Boucher, Z. Altamimi, Status of the realization of the BIH Terrestrial System. In: Babcock AK, Wilkins GA (eds) The Earth's rotation and reference frames for geodesy and geodynamics, Symposium no. 128, International Astronomical Union. Kluwer Academic, Dordrecht, p 107, 1986

W.H.M. Christie, *Results of the Astronomical Observations Made at the Royal Observatory Greenwich for the Year 1901* (His Majesty's Stationery Office, Edinburgh, 1903), p. xxxiii

M. Ekman, J. Ågren, *Reanalyzing Astronomical Coordinates of Old Fundamental Observatories using Satellite Positioning and Deflections of the Vertical*, Small Publications in Historical Geophysics 21 (Summer Institute for Historical Geophysics, Åland Islands, 2010)

Nautical Almanac Office, *Explanatory Supplement to the Astronomical Ephemeris and The American Ephemeris and Nautical Almanac* (Her Majesty's Stationery Office, London, 1961)

M. Feissel, Determination of the earth rotation parameters by the Bureau International de l'Heure, 1962–1979. Bull. Géod. **54**, 81–102 (1980)

I. Fischer, The figure of the earth – changes in concepts. Geophys. Surv. **2**(1), 3–54 (1975)

R. Forsberg, *A study of terrain reductions, density anomalies and geophysical inversion methods in gravity field modeling*, Department of Geodetic Science and Surveying Report No. 355 (Ohio State University, Columbus, 1984)

P.J.D. Gething, The collimation error of the Airy Transit Circle. Mon. Not. R. Astron. Soc. **114**, 415 (1954)

R. Grant, *History of Physical Astronomy, From the Earliest Ages to the Middle of the XIXth Century* (Henry G. Bohn, London, 1852), p. 461

B. Guinot, in *Irregularities of the polar motion*, ed. by D.D. McCarthy, J.D.H. Pilkington. Time and the Earth's Rotation, Proceedings of IAU Symposium, vol. 82 (D Reidel, Dordrect, 1979), p. 280

B. Guinot, in *History of the Bureau International de l'Heure*, ed. by S. Dick, D. McCarthy, B. Luzum. Polar Motion: Historical and Scientific Problems, Proceedings of IAU Colloquium 178, Astronomical Society of the Pacific Conference Series 208 (Astronomical Society of the Pacific, San Francisco, 2000), pp. 175–184

W.A. Heiskanen, H. Moritz, *Physical Geodesy* (W.H. Freeman, San Francisco, 1967)
C. Hirt, U. Marti, B. Bürki, W.B. Featherstone, Assessment of EGM2008 in Europe using astrogeodetic vertical deflections and omission error estimates from SRTM/DTM2006.0 residual terrain model data. J. Geophys. Res. **115**, B10404 (2010). doi:10.1029/2009JB007057
D. Howse, *Greenwich Time and the Longitude* (Oxford University Press, Oxford, 1997)
IAU, Resolution of Commission 19, in *Proceedings of the Thirteenth General Assembly, Prague 1967. Transactions of the IAU XIII B*, ed. By L. Perek (International Astronomical Union, 1968)
IERS (2010) IERS Technical Note 36, IERS Conventions (2010), IERS Conventions Centre, Verlag des Bundesamts fur Kartographie und Geodasie, Frankfurt am Main
IERS, IERS Annual Report 2013, in *International Earth Rotation and Reference Systems Service, Central Bureau*, ed. By W.R. Dick, D. Thaller (Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, 2014). Available at http://www.iers.org/IERS/EN/Publications/AnnualReports/AnnualReport2013.html?nn=94904
IUGG, Traveaux de l'I.A.G., XIII IUGG General Assembly, Resolutions, 1967
C. Jekeli, An analysis of vertical deflections derived from high-degree spherical harmonic models. J. Geod. **73**(1), 10–22 (1999). doi:10.1007/s001900050213
C. Jekeli, *Geometric Reference Systems in Geodesy* (Division of Geodesy and Geospatial Science, School of Earth Sciences, Ohio State University, Columbus, 2006)
F.G. Lemoine, S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, T.R. Olson, The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96, NASA Technical Publication TP-1998-206861, 1998
A.S. MacDonald, The Ordnance Survey and the Greenwich Meridian. Vistas Astron. **28**, 289–304 (1985)
S. Malys, J. Seago, N.K. Pavlis, P.K. Seidelmann, G.H. Kaplan, Why the Greenwich meridian moved. J. Geod. **89**(12), 1263–1272 (2015). doi:10.1007/s00190-015-0844-y
H. Moritz, I.I. Mueller, *Earth Rotation—Theory and Practice* (Ungar, New York, 1987)
I.I. Mueller, *Spherical and Practical Astronomy as Applied to Geodesy* (Ungar, New York, 1969)
W.H. Munk, G.J.F. MacDonald, *The Rotation of the Earth—A Geophysical Discussion* (Cambridge University Press, Cambridge, 1975), p. 61
W.A. Norton, *A Treatise on Astronomy, Spherical and Physical*, 4th edn. (Wiley, New York, 1872), p. 33
N.K. Pavlis, S.A. Holmes, S.C. Kenyon, J.K. Factor, An earth gravitational model to degree 2160: EGM2008. Presented at 2008 General Assembly of the European Geosciences Union, Vienna, Austria, 13–18 April 2008
N.K. Pavlis, S.A. Holmes, S.C. Kenyon, J.K. Factor, The development and evaluation of the Earth Gravitational Model 2008 (EGM2008). J. Geophys. Res. **117**, B04406 (2012). doi:10.1029/2011JB008916
A.R. Robbins, Time in geodetic astronomy. Surv. Rev. **19**(143), 11–16 (1967)
H.H. Schmid, *Three-Dimensional Triangulation With Satellites* (NOAA, Rockville, 1974), pp. 17–19
W. Torge, *Geodesy*, 3rd completely revised and extended edn. (Walter de Gruyter, Berlin, 2001)
C.C. Tscherning, Estimation of the longitude bias of the NWL9D coordinate system from deflections of the vertical, satellite altimetry and high degree spherical harmonic expansions. Bull. Géod. **60**, 29–36 (1986)
G.A. Wilkins, in *Project merit and the formation of the international earth rotation service*, ed. by S. Dick, D. McCarthy, B. Luzum. Polar Motion: Historical and Scientific Problems, Proceedings of IAU Colloquium 178. Astronomical Society of the Pacific Conference Series 208, 2000, pp. 187–200
Y.S. Yatskiv, Report of Commission 19: Rotation of the earth (Rotation de la Terre). Trans. Int. Astron. Union, Series A (Reports on Astronomy) **19A**, 193–205 (1985)

Chapter 29
Aspects of Time as It Relates to Space Geodesy

Ludwig Combrinck

Abstract High accuracy space geodesy relies on accurate and stable frequency standards, time interval counters, event timers and measured time offsets relative to UTC. These requirements are becoming more stringent as the global networks of space geodesy stations strive towards higher accuracy. This invariably leads to continual instrumentation upgrades and instrumental developments. In addition to the station frequency standard high stability and accuracy requirements, several relativistic considerations are required during data processing to reach optimal measurement accuracy. These aspects of time are discussed and quantified for each of the techniques of VLBI, GPS, SLR and LLR. It is proposed that fundamental space geodesy stations be equipped with the new generation of frequency standards that will allow chronometric geodesy. This will allow time determination within a stable and defined terrestrial reference frame, linked to the celestial reference frame via the Earth orientation parameters, and will also provide a link between the relativistic and classical geoid at these stations.

Keywords General and special theory of relativity • VLBI • SLR • LLR • GNSS • Chronometric geodesy

Introduction

There are four main space geodesy techniques; these all operate within large international networks to ensure global distribution and global participation. The techniques of Very Long Baseline Interferometry (VLBI), Global Navigation Satellite Systems (GNSS), Satellite and Lunar Laser Ranging (SLR or LLR) and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) all require very accurate time and frequency systems to operate at the required accuracy levels. Requirements between techniques vary, with VLBI requiring hydrogen maser frequency standards, whereas GNSS receivers can produce good results with a temperature compensated or ovenized crystal oscillator. In some

L. Combrinck (✉)
Hartebeesthoek Radio Astronomy Observatory, PO Box 443, Krugersdorp 1740, South Africa
e-mail: ludwig@hartrao.ac.za

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_29

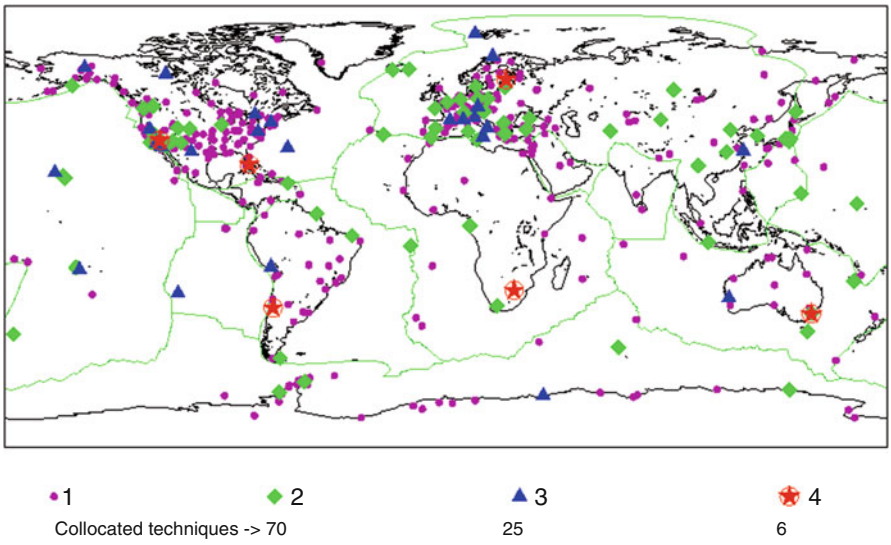


Fig. 29.1 Global map illustrating the distribution of space geodetic techniques. Very few sites are equipped with the four main techniques. The station in Chile has recently been relocated to La Plata (Argentina). It is clear that some areas are under-represented, e.g. the African continent in particular and large oceanic regions. Geodetic equipment distribution is dynamic as there is a continuous effort by the global community to improve and densify the networks (source: http://itrf.ensg.ign.fr/ITRF_solutions/2000/map.php)

cases, fundamental sites (geodetic observatories where four or more techniques are collocated) such as the Hartebeesthoek Radio Astronomy Observatory (HartRAO) will distribute the hydrogen maser 5 MHz standard frequency to all instruments, including SLR stations and GNSS receivers to ensure optimal accuracy for all systems. Figure 29.1 illustrates the current global network of geodetic techniques.

These different space geodetic techniques all operate within the International Association of Geodesy (IAG) (http://www.iag-aig.org/) services; links to these services can be found at the IAG website. The IAG currently drives a global project (Global Geodetic Observing System, GGOS) which, in a seamless way, throughout the services, aims to provide modernized geodetic infrastructure to meet the high accuracy and stable reference frame and other products required to meet the stringent requirements to monitor Earth in a holistic manner. Although the classical three main pillars of geodesy (Earth's shape, global gravity field and the Earth's rotational motion) are primarily the observables, the modern theory of gravity (space-time curvature) as described by Einstein's general theory of relativity and even the special theory of relativity play such a large role in modern space geodesy that space-time curvature has become the fourth observable/pillar of space geodesy. Time itself becomes an observable in this context. Proper time of a GNSS satellite is not the same as proper time for a GNSS receiver located on the geoid; a GNSS receiver that is not located on the geoid is slightly removed from terrestrial time (TT), as TT is time defined on the geoid (equipotential surface). This fourth pillar

will become more relevant as the techniques of space geodesy move into the planetary system and eventually into outer space beyond the planetary system.

With this continuous growth in the networks and continuous improvement in the techniques, there is a commensurable growth in applications of the products of space geodesy and a continuous requirement for an increase in accuracy and stability of the derived geodetic reference frames. The objectives of GGOS are to reach and maintain global positional accuracy of 1 mm and to create a terrestrial reference frame that is stable to 0.1 mm. These are very stringent requirements by any standards. As the space geodetic observables are dependent to a large extent on stable frequency and time, requirements of time and frequency will no doubt also increase.

This is particularly true for the new technique of chronometric geodesy, which is still in its infancy, but directly requires clock stabilities at the fractional frequency error level of 1×10^{-18}. A typical hydrogen maser clock currently has a comparative frequency stability of 2×10^{-15}. Chronometric geodesy directly observes gravity potential changes through oscillator frequency changes (first predicted through the general theory of relativity); an oscillator that is stable to 1×10^{-18} will be able to detect 1 cm changes in height. When these "super clocks" become available commercially, it will be necessary to equip fundamental geodetic sites with at least one; this will be crucial in locating the "relativistic geoid" (surface closest to mean sea level on which precise clocks have the same rate) and matching it with the "classical geoid" (equipotential surface). Terrestrial time (TT) is defined as the proper time of a clock fixed to the geoid; however, the clock rate changes when the clock is moved above or below the geoid (due to gravitational time dilation). In practice this effect is currently mainly ignored, but with the advent of super clocks, this will no longer be the case.

Such clocks will lead to interesting (and more accurate) determinations of the geoid, its undulations and variations due to gravitational potential changes at a given location. In effect, time will be linked to gravitational potential in a practical way and may well lead to a redefinition of "terrestrial time" to include a clock's position in a reference frame where its coordinates (x, y, z) and the gravitational potential (U) (and its variations) are inextricably interlinked practically; this will mean that clocks in effect will provide time *and* gravitational potential and the rate of time at a particular position.

The HartRAO site can be seen in Fig. 29.2. Two radio telescopes used for space geodesy and astronomy are visible, the 26 m radio telescope (used for celestial reference frame measurements) in front and 15 m radio telescope (used for Earth Orientation Parameters) in the rear. Towards the centre and right can be seen an SLR station (Moblas-6, part of the NASA SLR network) and the control centre and run-off shelter for an LLR system being developed in collaboration with NASA and the Observatoire de la Cote d'Azur (OCA) of France. In the near foreground, two vents of an underground vault housing a gravimeter and seismometer can be seen, part of a network we are installing in collaboration with several universities. In the near foreground is a GNSS antenna of the HartRAO GNSS receiver (HRAO), part of the International GNSS Service (IGS). A 5 MHz timing signal is fed via optical fibre to the SLR, LLR and GNSS receivers.

Fig. 29.2 The HartRAO site, several instrumentation used in space geodetic techniques including geophysical instruments are collocated on-site

The hydrogen masers are located in the main building seen in the left of Fig. 29.2. Two of the three hydrogen masers of HartRAO can be seen in Fig. 29.3 and the frequency distribution rack in Fig. 29.4. It is important to have an independent backup power supply to ensure that the masers (and the maser room's thermal control circuitry) remain powered-up during power grid outages, otherwise VLBI and other instruments will be adversely affected as it may take several days for the maser to regain its prior stability. To monitor hydrogen maser frequency stability and to keep the clocks within a certain offset from UTC (~2 microseconds), a comparison is made between UTC as provided by a GPS receiver (1 PPS). This can be done to the 20 nanosecond RMS level and provides information on maser clock ageing and frequency offset. Not one hydrogen maser is the same; therefore they are all independently monitored and their characteristics logged.

Very Long Baseline Interferometry (VLBI)

The technique of VLBI requires a number of radio telescopes (globally distributed) appropriately equipped with receivers at S and X band, as well as a hydrogen maser as frequency reference. A new global network of VLBI radio telescopes

Fig. 29.3 Hydrogen masers at HartRAO; these are located in a thermally stabilized room to reduce thermal fluctuations. (**a**) Model EFOS-28, installed November 2003, which is an EFOS-C model. (**b**) HartRAO's first maser (EFOS-6) acquired for VLBI purposes; installed May 1985, this is an EFOS-A model. Not shown is iMaser-72, installed April 2012, which is an iMaser 3000 model. Room temperature is controlled to within ±2 °C

(~13.2 m diameter, slewing rate of 12° in azimuth and 6° in elevation) is being developed; this is being established to meet the requirements of the Global Geodetic Observing System (GGOS) and is known as VGOS telescopes (VLBI Global Observing System); these will operate in the 2–14 GHz frequency range and require wide-band receivers and feeds. The legacy (larger, slower, S and X band) antennas will continue to operate at some level within the VGOS network.

Hydrogen maser clocks are a requirement for VLBI (Wei-qun et al. 2001), and currently other types of frequency standards (rubidium, caesium) are not stable enough. Hydrogen maser clocks have a typical clock frequency accuracy of $\pm 5 \times 10^{-13}$ and a stability of 9.7×10^{-15} when measured over a 24 h period. A stability of 9.7×10^{-15} indicates that there are 9.7 excess counts for every 1×10^{15} counts or pulses in a 24 h period; the clock stability and accuracy are thus normally expressed in parts per so many counts. Signal phase coherence is an important requirement of the VLBI technique.

There are some limitations on signal coherence due to the ionosphere (dispersive for radio waves) and equipment-related issues (spectral purity and stability) such as

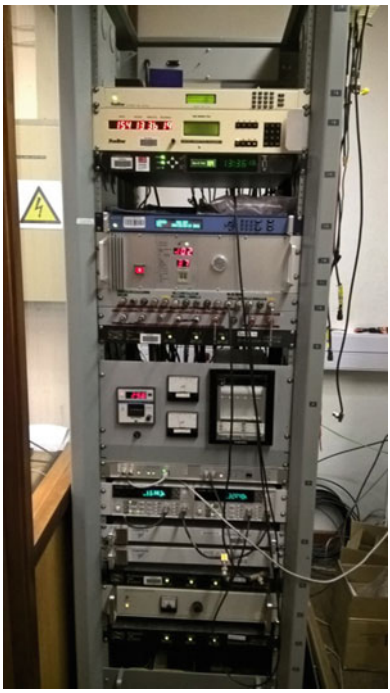


Fig. 29.4 Time and frequency distribution rack. A number of geodetic and astronomical equipment rely on a 5 MHz reference distributed from the hydrogen masers. Photograph courtesy of HartRAO

local oscillators and frequency multipliers. The VLBI observable is the group delay T_{group} given by

$$T_{\text{group}} = \partial\phi/\partial\omega, \tag{29.1}$$

where the group delay (Eq. (29.1)) is the first-order derivative of the phase (ϕ) delay. The group delay has a number of components as seen in Eq. (29.2):

$$T_{\text{group}} = T_{\text{g}} + T_{\text{inst}} + T_{\text{trop}} + T_{\text{ion}} + T_{\text{rel}} + \cdots \tag{29.2}$$

Here T_{g} is the geometric delay, T_{inst} represents delay in the instrumentation (signal path), delays due to the troposphere and ionosphere are represented by T_{trop} and T_{ion}, respectively, and T_{rel} is the delay due to relativistic effects. To interpret the phase, one has to resolve the inherent ambiguity (i.e. determine the number of phase cycles). To estimate the number of cycles, it is essential to have an accurate geometric model (<20 picosecond or 5 mm at 8 GHz). This is in general not available and therefore τ_{φ} cannot be used directly.

The coherence (where signals maintain a fixed phase relationship with each other) time of a VLBI system is determined by the frequency standard, which normally acts as the fundamental frequency source for the system's equipment (oscillators, time code generators, etc.). The contribution of clock phase stability can be seen in Eq. (29.3) which expresses signal-to-noise ratio of the VLBI signal:

$$\mathrm{SNR} = \frac{SA\sqrt{2B\tau_c}}{kT_s}. \tag{29.3}$$

Within Eq. (29.3), τ_c is the coherent integration time, S is the source density flux, A represents the geometric mean

$$\bar{A}_{\mathrm{geom}} = \sqrt[n]{\prod_{i=1}^{n} A_i} = \sqrt[n]{A_1 \cdot A_2 \cdot \ldots \cdot A_n}, \tag{29.4}$$

of the collecting areas of the telescopes, the receiver system bandwidth is given by B, k is Boltzmann's constant and the geometric mean of the system temperatures of the radio telescopes is represented by T_s (Moran 1989). The approximate coherent integration time τ_c is contained within Eq. (29.5),

$$\omega \times \tau_c \times \sigma_y(T) = 1. \tag{29.5}$$

Here $\sigma_y(T)$ is the two-sample Allan variance and ω the local oscillator frequency (radians per second). The Allan variance is expressed as

$$\sigma_y^2(\tau) = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} \left(y_{i+1} - y_i\right)^2. \tag{29.6}$$

In Eq. (29.6), m is the number of samples, y_i the value of sample i, y_{i+1} is the value of the $i+1$ sample, the sample time is represented by τ, and $\sigma_y^2(\tau)$ represents the Allan variance.

A simple example will illustrate the importance of reference frequency stability requirements. To achieve signal coherence for a VLBI observation period of 1000 s, with local oscillator frequency set to 8.0 GHz, the two (or more) frequency standards within the interferometry network need to maintain relative stability of $\sim 2 \times 10^{-14}$. Requirements of the VLBI technique demand that the influence of clock stability extends throughout the VLBI hardware.

Satellite Laser Ranging (SLR)

Satellite laser ranging (SLR) is a geodetic technique which utilizes an optical telescope equipped with a suitable laser and photon detection system. A short-pulse (<200 ps) laser transmits though an exit tube (typically a refractor) fixed parallel to and co-aligned with the telescope. Stable and accurate timing equipment is required (the timing clock is typically based on 5 or 10 MHz cesium, rubidium or hydrogen maser frequency standard) as well as a host of peripheral equipment, which leads to an integrated system which allows transmitting and receiving laser

pulses to satellites. Only satellites equipped with corner cube reflectors are used; typical applications are precise orbit determination (e.g. to determine orbits of GNSS satellites), useful for space oceanography (calibrating radar equipped satellites) and determination of an absolute scale factor for orbit determination, through using LAGEOS ranging data to determine the gravitational constant (Dunn et al. 1999).

The basic components of an SLR system are described in Fig. 29.5. An important subsystem is the frequency standard; this unit feeds the time interval counter (used to measure the time of flight (TOF) of a laser pulse). It also provides time to the control computer and telescope servo control. The frequency standard and distribution of frequency throughout the other subsystems of the SLR are extremely important so as to have all equipment synchronized. This can be illustrated with an example; the velocity of the satellite LAGEOS within a 1-day (arbitrarily selected) arc (~6.3 orbital revolutions) had a minimum velocity of 5645 m/s and a maximum velocity of 5806 m/s. This would require a timing precision of 1.7×10^{-7} or 0.17 μs to register the epoch of the observation (this accuracy is of course easily attainable with the clocks currently used). If this observation accuracy is not good enough (clock error), a small time bias will develop as the epoch of the observation is either delayed or advanced. In general, time bias values of SLR stations are at the few

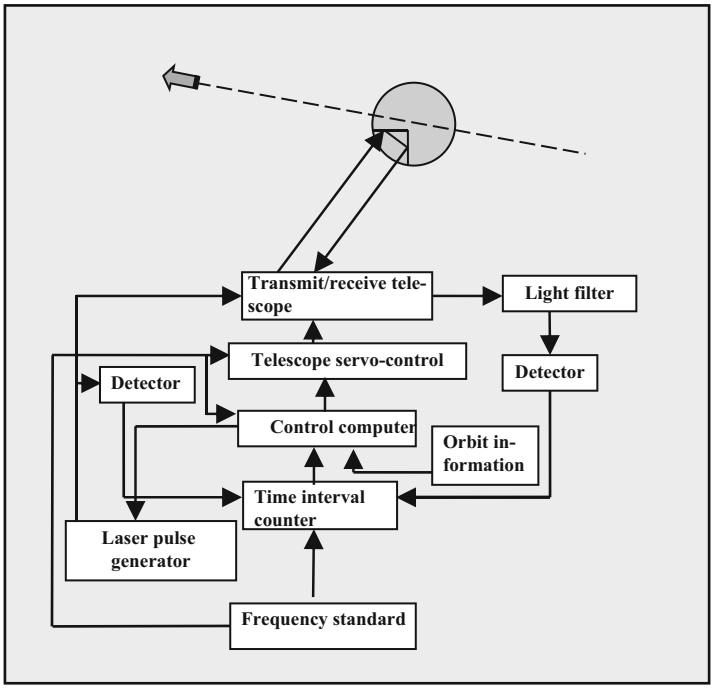


Fig. 29.5 Basic satellite laser ranging system, describing components and signal paths (Combrinck 2010)

microsecond levels, normally determined during post-analysis of the tracking data and corrected during the analysis; typically the clock is not the limiting factor, but there are delays in the system electronics and software. Considering the satellite range, the velocity of light (~3×10^8 m/s) leads to a two-way range precision of 0.15 mm/ps; therefore to achieve a few millimetres precision level, a precision of at least 20 ps should be attained in the timing system and electronics of the interval counter or event timer.

Global Navigation Satellite Systems (GNSS)

The Global Navigation Satellite Systems (GNSS) comprise the various space-based satellite navigation systems. These systems are continuously evolving; currently GNSS is made up of the Global Positioning System (GPS) of the USA, the Beidou system (PRC) and the GLONASS (Russian Federation). Other systems are also at various stages of development such as the European Union's Galileo system. Here we will use GPS as an example; the principles involved are valid for all GNSS. In particular we will discuss the effect of relativity on GPS, but in general GPS should just be seen as part of GNSS. There are certain relativistic effects that have to be taken into account when reducing VLBI and SLR data as well; these are discussed in detail in Schuh and Böhm (2013) and Combrinck (2010), respectively. The orbital height of GPS satellites is ~20,200 km, and orbital periods are ~718 min and velocities ~3874 m/s. The satellite's velocity relative to the geocentre, the rotation of Earth and the gravitational potential difference between the GPS receiver and satellite requires that GRT needs to be considered in timing (frequency) and orbital parameters. Geocentric Coordinate Time (*TCG*) is the basic theoretical timescale for geodesy and geophysics and is the coordinate time of the Geocentric Celestial Reference System (GCRS), which has coordinates $\left(T, \vec{X}\right)$ (Müller et al. 2008; Petit and Luzum 2010).

Time dilation and gravitational redshift affects the clocks of GNSS satellites. The relative motion between the GNSS receiver and satellite results in special relativistic time dilation (due to its velocity the GNSS clock rate is slower relative to the GNSS receiver located on Earth), and the differences in the gravitational potential present at the GNSS receiver and GNSS satellite result in gravitational redshift. The nett effect is that the effect due to the lower gravity potential is larger and the GNSS clock rate is faster; this implies that the clock located on the satellite must be set to run slower to have its proper time match that of the GNSS receiver on Earth; this is a practical proof of GRT. Therefore, due to these two relativistic effects, and to approximately align the GPS clocks with terrestrial time (TT), onboard oscillators must be adjusted slightly (lower) in frequency. Other second-order effects resulting from the elliptical (non-circular) orbits have to be considered at the receiver during observational data processing by introducing an adjustment of the order of $2\left(\vec{r} \cdot \dot{\vec{r}}\right)/c^2$, where the position vector of the satellite is denoted by $\vec{r}$ (Senior et al. 2008).

Conclusion

Time and frequency are of extreme importance in space geodetic measurements; without stable and accurate clocks, all space geodesy techniques would be adversely affected. Satellite laser ranging requires measurement of laser pulse time-of-flight measurements at the picosecond level, VLBI requires hydrogen maser atomic standards and GNSS receivers require at least temperature compensated or ovenized crystal oscillators to achieve good accuracy and several relativistic considerations. With the advent of chronometric geodesy, a new tool for geodesy is in its infancy, and we are bound to have many exciting applications where clock frequency and clock rate become tied to position and gravitational potential.

References

L. Combrinck, Satellite laser ranging, in *Sciences of Geodesy I, Advances and Future Directions*, ed. by G. Xu (Springer, Germany, 2010), pp. 301–336

P. Dunn, M. Torrence, R. Kolenkiewicz, D. Smith, Earth scale defined by modern satellite ranging observations. Geophys. Res. Lett. **26**(10), 1489–1492 (1999)

J. Moran, Introduction to VLBI, in *Very Long Baseline Interferometry, Techniques and Applications*, ed. by M. Felli, R.E. Spencer (Kluwer Academic Publishers, 1989), pp. 27–45

J. Müller, M. Soffel, S.A. Klioner, Geodesy and relativity. J. Geod. **82**, 133–145 (2008)

G. Petit, B. Luzum, in *IERS Conventions (2010)*, ed. by G. Petit, B. Luzum (IERS Technical Note; No. 36), Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, 2010. Available at: http://www.iers.org

H. Schuh, J. Böhm, Very long baseline interferometry for geodesy and astrometry, in *Sciences of Geodesy II*, ed. by G. Xu (2013), pp. 339–376

K.L. Senior, J.R. Ray, R.L. Beard, Characterization of periodic variations in the GPS satellite clocks. GPS Solut. **12**, 211–225 (2008)

Z. Wei-qun, L. Chuan-fu, Y. Shuang-lin, W. Guan-zhong, Z. Yi-ping, Y. Pei-hong, Z. Jian, A study and performance evaluation of hydrogen maser used in Chinese mobile VLBI stations. Chin. Astron. Astrophys. **25**(3), 390–397 (2001)

Chapter 30
Pulsars: Celestial Clocks

R.N. Manchester, L. Guo, G. Hobbs, and W.A. Coles

Abstract Pulsars are rapidly rotating neutron stars, with most of the known examples located within our Milky Way Galaxy. The class of "millisecond pulsars" (MSPs) have remarkably stable pulse periods, with a stability over long intervals comparable to that of the best atomic frequency standards. Timing observations of such pulsars in binary systems with another star have been used to make the most stringent tests of general relativity in strong gravitational fields. Observations of many MSPs, widely distributed across the sky and forming a "pulsar timing array" (PTA), have been used to set limits on the strength of the gravitational-wave background at nanohertz frequencies. These limits are beginning to constrain current ideas about the formation and evolution of supermassive binary black holes in the cores of distant galaxies. PTA observations can also be used to define a "pulsar timescale" that can limit or measure instabilities in terrestrial timescales over intervals of years and decades. Pulsar timescales form a useful "secondary standard" that is totally independent of terrestrial time and frequency standards and will be continuous for billions of years.

Keywords Time—pulsars • general

Introduction

Pulsars are rotating neutron stars that send out beams of emission that we see as pulses as they sweep across the Earth. The vast majority of the ~2500 known pulsars[1] are located within our Milky Way Galaxy, mostly at distances of several thousand light

[1]For an up-to-date list of the currently known pulsars and their basic properties, see http://www.atnf.csiro.au/research/pulsar/psrcat and Ref. (Manchester et al. 2005).

R.N. Manchester (✉) • G. Hobbs
CSIRO Astronomy and Space Science, Sydney, NSW, Australia
e-mail: dick.manchester@csiro.au

L. Guo
Shanghai Astronomical Observatory, CAS, Shanghai, China

W.A. Coles
Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA, USA

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_30

years from the Sun. Pulsars fall into two main groups: "normal" pulsars with pulse (rotation) periods P of ~30 ms up to ~12 s and "millisecond pulsars" (MSPs) with pulse periods between 1.3 and 50 ms. Normal pulsars are relatively young with "characteristic ages" $\tau_c = P/\left(2\dot{P}\right)$ (where $\dot{P}$ is the first time derivative of P) typically ~10^6 year, whereas MSPs are much older with τ_c typically ~10^9 year. Most MSPs are in a binary system with another star, whereas most normal pulsars are single. The high fraction of binary MSPs is consistent with the idea that they obtain their high rotation speed by accreting matter from an evolving binary companion in an X-ray-emitting phase (Bhattacharya and van den Heuvel 1991). This has the effect of "recycling" a long-period and probably dead neutron star to a new life as an MSP. In some way that is not fully understood, the accretion process also greatly reduces the pulsar magnetic field, resulting in a reduction of $\dot{P}$ by about five orders of magnitude and hence the large characteristic age. Normal pulsars do exhibit unpredictable fluctuations in their period, but the strength of these fluctuations is related to $\dot{P}$ (Shannon and Cordes 2010). With their low $\dot{P}$, MSP periods are extremely stable.

This great period stability makes pulsars extremely useful tools for a variety of applications. Foremost among these is using pulsars as probes of gravitation, firstly, using binary pulsars as tests of gravitational theories and, secondly, using observations of a large sample of MSPs as a "pulsar timing array" (PTA) to detect low-frequency (nanohertz) gravitational waves. However, there are numerous other applications, including investigating the interior structure of neutron stars, studies of the properties of the ionized interstellar medium between us and the pulsar, measuring the trajectory of the Earth with respect to the solar system barycentre and, importantly in the present context, establishing an independent long-term standard of time, a "pulsar timescale". Tests of gravitational theories using binary pulsars are described in section "Tests of Gravitational Theories Using Binary Pulsars", pulsar timing arrays in section "Pulsar Timing Arrays" and pulsar timescales in section "Pulsar Timescales". In this work, we make use of a data set created as part of the International Pulsar Timing Array project (Hobbs et al. 2010).

Tests of Gravitational Theories Using Binary Pulsars

The first-known binary pulsar, PSR B1913+16, was discovered at Arecibo Observatory by Hulse and Taylor (1975). This system has an orbital period of just 7.75 h, but most remarkably, the pulsar has an orbital velocity v that peaks at about 320 km s^{-1} or ~0.1% of the velocity of light! Since pulsar periods can be measured with precisions exceeding 1:10^{12}, this means that relativistic effects that vary as $(v/c)^2$ are easily measurable in this system. The first such effect detected was the relativistic precession of the longitude of periastron, $\dot{\omega}$, which had the amazingly high value of ~4.2° per year. It is interesting to note that this is the same effect as the relativistic perihelion advance of Mercury that was used by Einstein in 1915 as an

observational verification of general relativity, but this had a value of just 43 arcseconds per century! The second relativistic effect detected was the combination of variations in second-order Doppler shift and gravitational redshift as the pulsar moved around its eccentric orbit. Predictions of post-Keplerian orbital parameters by different theories of relativistic gravity have different dependencies on the binary parameters, but in general relativity (GR), these two effects depend on just the accurately measured binary Keplerian parameters and the pulsar mass and the companion mass. The two effects have different dependencies on the masses, and measurement of them (Taylor et al. 1979) gave estimates of the masses which were about 1.44 M_{Sun} and 1.39 M_{Sun}, respectively, consistent with the idea that both the pulsar and its companion are neutron stars.

Given the two masses, it is possible to predict the value of other relativistic effects. Specifically, GR predicts that the binary orbit should decay as a result of losing energy to gravitational waves. The early measurements (Taylor et al. 1979) hinted at this effect and recent measurements (Weisberg et al. 2010) have confirmed it to high precision as shown in Fig. 30.1, with the ratio of the measured to

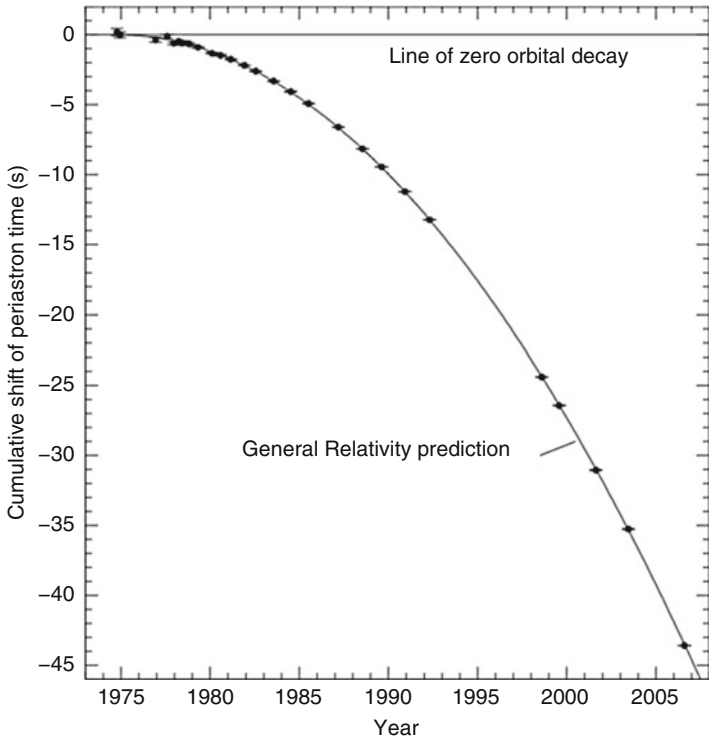


Fig. 30.1 Decay of the orbit of PSR B1913+16, measured as a progressive shift in the time of periastron passage relative to a constant orbital period, and the prediction of GR for this effect based on energy loss from the system in the form of gravitational waves (Weisberg et al. 2010)

predicted values being 0.997 ± 0.002. This measurement confirmed that GR gave an accurate description of the motions of the pulsar and its companion and provided the first observational evidence for the existence of gravitational waves.[2] The discovery of this system and the gravitational tests that it made possible were recognized by the award of the Nobel Prize in Physics to Joseph Taylor and Russell Hulse in 1993.

The Double Pulsar

An even more remarkable double neutron star binary system, the "Double Pulsar", PSR J0737−3039A/B, was discovered at Parkes in Australia in 2003 (Burgay et al. 2003; Lyne et al. 2004). This system had an orbital period of just 2.45 h and was (and remains) unique in that both neutron stars were observed as pulsars.[3] One is a recycled MSP with a pulse period of 22 ms, and the other is a much younger but slower pulsar with a pulse period of 2.7 s. This system is even more relativistic than the Hulse-Taylor binary with a predicted (and observed) periastron advance of 16.9° per year, more than four times that of PSR B1913+16.

Based on 20 years of timing with the Parkes and Green Bank telescopes, a total of six relativistic parameters have now been measured for this system. These, together with the mass ratio obtained from the ratio of orbit semi-major axes for the two pulsars, give five independent tests of GR (cf. Kramer et al. 2006) as illustrated in Fig. 30.2. As the inset shows, all constraints are consistent with a tiny region near the apex of the white region, accurately defining the masses of the two stars. Overall, these results confirm the accuracy of GR at the 0.02% level, an order of magnitude more precise than the PSR B1913+16 result. While the strongest constraint comes from the Shapiro-delay measurement of sin$i$ (where i is the inclination angle of the orbit normal), the observed orbit decay is fully consistent with GR's prediction of gravitational-wave emission and is now more precise than the PSR B1913+16 result. It is remarkable that, despite the invention of many alternative theories of gravity over 100 years since Einstein created general relativity, it seems as though he got it right the very first time!

[2]The detection of binary orbit decay due to the emission of gravitational waves is generally termed an "indirect" detection of gravitational waves in contrast to the recent direct detection of a gravitational-wave burst by LIGO (Abbott et al. 2016).

[3]The B pulsar disappeared from view in 2008, most likely because of spin-axis precession, but it is expected to return to view sometime in the next 20 years (Perera et al. 2010).

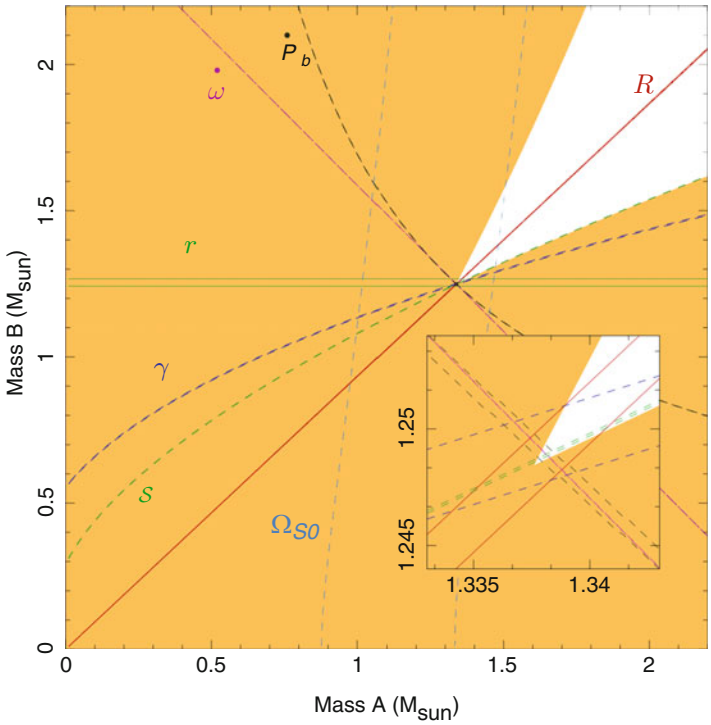


Fig. 30.2 Plot of the mass of PSR J0737−3039B versus the mass of PSR J0737−3039A with the various constraints from Newtonian physics and GR marked. All lines are double (although this is not obvious for most parameters) with the separation of the lines representing the uncertainty in the constraint. The various constraints are mass ratio from the Keplerian orbit sizes R (*red line*), the periastron advance $\dot{\omega}$ (*magenta dashes*), the time dilation/second-order Doppler term γ (*blue dashes*), orbit decay $\dot{P}_b$ (*black dashes*), Shapiro sin i (*green dashes*), Shapiro range r (*green lines*) and geodetic precession Ω_{SO} (*grey dashes*). The *yellow-shaded* regions are disallowed by the observed Keplerian parameters and the requirement that sin $i \leq 1$. The inset shows a magnified version of the central region. [Kramer et al. in preparation]

Pulsar Timing Arrays

A "pulsar timing array" (PTA) consists of many pulsars widely distributed across the sky with frequent high-precision timing observations over a long data span (Foster and Backer 1990; Hellings and Downs 1983). The principal scientific objective for PTAs is the detection of low-frequency gravitational waves from binary supermassive black holes in the cores of distant galaxies, but there are many secondary objectives, including the establishment of a "pulsar timescale". Only MSPs have sufficient timing stability to be useful for PTAs.

There are three main PTAs currently operating around the world: the European Pulsar Timing Array (EPTA), the North American pulsar timing array (NANOGrav) and the Australian Parkes Pulsar Timing Array (PPTA). Each of them is regularly timing 20–40 MSPs with data spans of up to 20 years (Arzoumanian et al. 2015;

Desvignes et al. 2016; Manchester et al. 2013). For the best-timed pulsars, uncertainties in pulse times of arrival (ToAs) are less than 100 ns. The three PTAs have formed a collaborative partnership known as the International Pulsar Timing Array (IPTA) (Hobbs et al. 2010; Manchester 2013) with the objective of enhancing the sensitivity for PTA science objectives by combining the data sets of the participating PTAs (Verbiest et al. 2016). Overall, about 70 MSPs are being regularly timed with ~20 being timed by more than one PTA.

Detection of gravitational waves (GWs) with a PTA relies on the fact that passage of a GW across the Earth results in correlated timing residuals[4] in different pulsars. Other effects also result in correlated residuals, for example, irregularities in the reference timescale used to time the pulsars or errors in the solar system ephemeris used to transfer the observed ToAs to the solar system barycentre, a (nearly) inertial reference frame. These different effects can be separated by their different dependence on pulsar sky position. For example, clock errors affect all pulsars equally, giving a spatial monopole signal, whereas solar system ephemeris errors, which effectively give an incorrect position for the Earth, result in a dipole spatial signature. In contrast, GWs give a quadrupolar signature in the observed ToA residuals (Hellings and Downs 1983) which (in principle) can be identified.

Up to now, no positive detection of nanohertz GWs has been made by PTAs. The best current limit comes from PPTA observations at 10-cm wavelength (3 GHz) of four well-timed pulsars (Shannon et al. 2015). The derived limit is $A_{1\mathrm{year}} < 1.0 \times 10^{-15}$, where $A_{1\mathrm{year}}$ is the characteristic GW strain at a GW frequency of 1 year^{-1}, defined by $h_c = A_{1\mathrm{year}} f^{-2/3}$, where f is the GW frequency in units of year^{-1} (Phinney 2001). This limit essentially rules out all current standard models for the formation and evolution of binary supermassive black holes in distant galaxies. It is possible that the late evolution of these binary systems is dominated by energy loss to surrounding stars or gas rather than emission of gravitational waves (Kocsis and Sesana 2011; Ravi et al. 2014), but there are other possible explanations for the non-detection (Sesana et al. 2016).

Further analysis of the highest-quality PPTA data set, that for PSR J1909–3744 at 10-cm wavelength, has revealed a puzzling dependence on the solar system ephemeris used in the data analysis. Figure 30.3 shows the timing residuals obtained with the JPL ephemerides DE418 (Folkner 2007) and DE435 (Folkner et al. 2016). In the top subpanel, the pulsar timing residuals (analysed with DE418) are nearly flat ("white") which suggests that the timing model is accurate and that the residuals are dominated by random noise. In contrast residuals obtained by simulating timing residuals with the same sampling representing the differences between DE418 and DE435 have correlated (or "red") offsets from zero. In the bottom panel, we show the pulsar observations analysed with DE435 and simulated DE435-DE418 points

[4]Timing residuals are the difference between observed and predicted ToAs, where the predictions are based on a model for the pulsar, including its astrometric parameters (position, proper motion, etc.) and timing parameters (pulse period, slow-down rate, dispersion measure and binary parameters if applicable) as well as the solar system ephemeris used in the analysis.

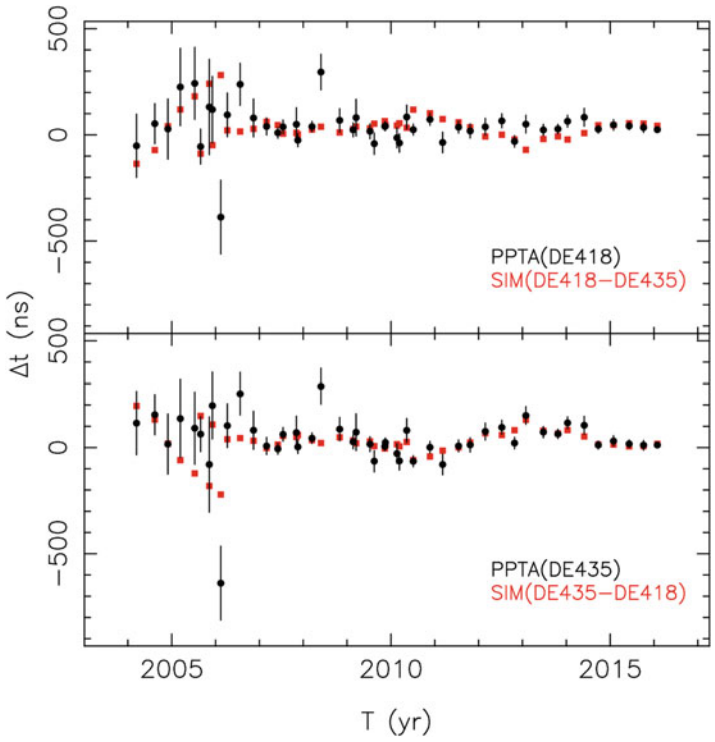


Fig. 30.3 Timing residuals for PSR J1909−3744 measured at Parkes at 10-cm wavelength (*black points*) compared to simulated timing residuals representing the differences between JPL epemerides DE418 and DE435 (*red points*). In the *top panel*, the pulsar data are analysed using the DE418 ephemeris, whereas in the *bottom panel*, the DE435 ephemeris was used. [Shannon et al. in preparation]

(i.e. the inverse of the points plotted in the top panel). It is clear that the pulsar residuals now have essentially identical red noise to the DE435-DE418 points. Unless there is an unlikely conspiracy such that PSR J1909−3744 has the same intrinsic red timing noise as the difference between the ephemerides, these results suggest that DE418 gives a better description of the position of the Earth relative to the solar system barycentre than does DE435. This is surprising since DE435 has the benefit of nearly a decade of extra observations which one would expect to make it more accurate. Further investigation of this problem is clearly needed.

Pulsar Timescales

Pulsar ToAs are measured relative to a local clock and then transferred to a reference timescale, normally TT(BIPMxx) (Petit 2007). As discussed above, PTAs allow irregularities in the reference timescale to be isolated from other intrinsic or extrinsic irregularities since they are identical for all pulsars in the PTA. The pulsar timing

analysis program TEMPO2 (Hobbs et al. 2006) enables extraction of the "common mode" or clock term from the residuals of all pulsars in the PTA. This clock term represents deviations of "pulsar time" from the reference timescale. However, one can view the problem from the opposite direction, defining pulsar time to be the reference timescale and measuring deviations of the atomic timescale from it. One important proviso on this technique is that linear and quadratic terms in the clock term are absorbed into the pulsar parameters since these terms are a priori unknown. Therefore, only higher-order terms in the differences between pulsar time and atomic time can be measured. Another important point is that the pulsar timescale is totally independent of the original reference timescale. This technique simply measures the differences between the reference timescale and pulsar time.

In 2012, Hobbs et al. (2012) analysed the PPTA DR1e data set (Manchester et al. 2013) to obtain the first pulsar timescale, TT(PPTA11), having a precision comparable to that of international reference atomic timescales. As an example, they chose to use TT(TAI) as the reference timescale and then compared the derived clock term with the known difference between TT(BIPM11) and TT(TAI). The derived clock term was consistent with the difference between these two reference timescales. In Fig. 30.4, we show a reanalysis of this data set using improved noise models for the pulsars (Reardon et al. 2016). It is clear that pulsar time tracks the

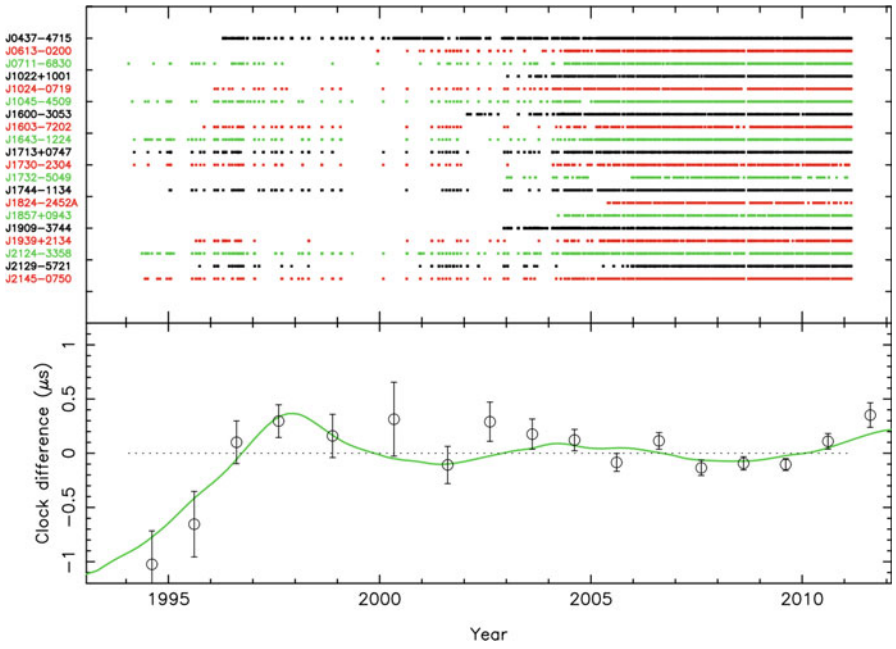


Fig. 30.4 *Upper panel*: ToA sampling for the 20 PPTA pulsars used in this analysis. *Lower panel*: The points show the measured differences and their uncertainties between pulsar time and the reference timescale TT(TAI), and the line shows the differences between TT(BIPM11) and TT(TAI) with linear and quadratic terms subtracted

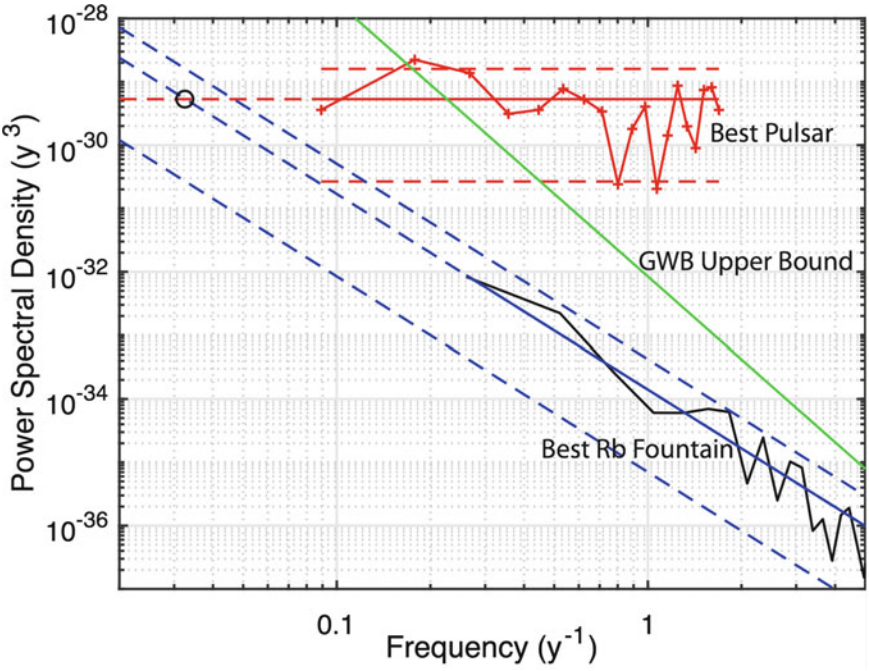


Fig. 30.5 Power spectra of the irregularities in timing residuals for PSR J1909−3744 based on Parkes 10 cm timing observations and time offsets between the best of four Rb fountain clocks at USNO relative to the average of all four fountains (Matsakis, private communication, 2016). The parallel dashed lines show the 90% confidence limits on each spectrum. The green sloping line corresponds to a GWB limit $A_{1\ \mathrm{year}}$ of 10^{-15} (Shannon et al. 2015)

known differences between TT(BIPM11) and TT(TAI) accurately within the uncertainties. This can be taken as an independent verification that the methods used to derive TT(BIPM11) from TT(TAI) are correct.

Since ToA uncertainties are typically several hundred nanoseconds, pulsar time cannot track short-term variations in the atomic timescale. It takes years or decades for the frequency precision of the pulsar timescale to reach values comparable to those of atomic timescales. This is illustrated in Fig. 30.5 which shows power spectra for irregularities in the residuals of our best-performing pulsar, PSR J1909−3744 at 10-cm wavelength (Shannon et al. 2015), and one of the best-performing atomic frequency standards over long time intervals, a rubidium fountain at USNO (Peil et al. 2014). The pulsar residual power spectrum is white with an rms timing residual of 99 ns. In contrast, the irregularity spectrum for the Rb clock is very red with a power-law slope of about −3. At high modulation frequencies, the atomic clock is orders of magnitude more stable than the pulsar, but extrapolation of both spectra suggests that at a frequency of about 0.035 cycles year^{-1}, the two timescales will have comparable stability. At timescales in excess of 30 years, it is likely that the stability of the pulsar clock will exceed that of this Rb fountain.

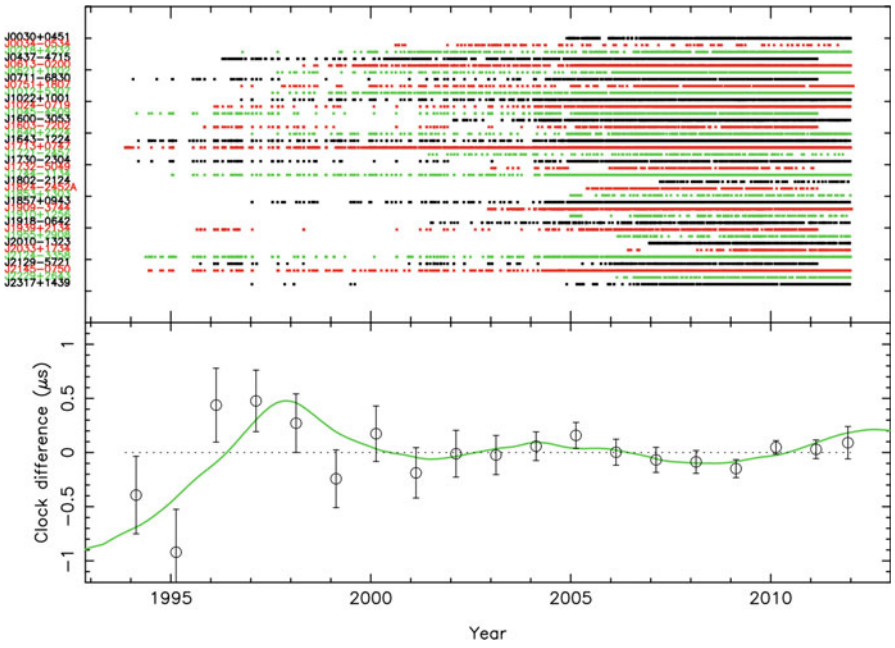


Fig. 30.6 A realization of a pulsar timescale based on the 37 most significant pulsars in the IPTA DR1 data set as shown in the *upper panel*. As in Fig. 30.4, the points show the measured differences and their uncertainties between pulsar time and the reference timescale TT(TAI), and the line shows the differences between TT(BIPM11) and TT(TAI) with linear and quadratic terms subtracted

This analysis is based on observations of just a single pulsar, albeit our most stable one. Greater stability can be achieved by combining data from many pulsars in a PTA. The IPTA offers the best data set for this (and other PTA objectives) since it has a greater number of precisely timed pulsars, longer data spans and more frequency diversity enabling better correction for interstellar medium propagation effects (Verbiest et al. 2016). In Fig. 30.6 we show the results of a preliminary analysis of the current best IPTA data set (Guo et al. in preparation). While only marginally better than the best PPTA result shown in Fig. 30.4, analysis of an improved and extended IPTA data set has the potential to provide the most precise pulsar timescale achievable at this time. In the future, as current PTA data sets become longer, and new and larger PTA data sets from large radio telescopes such as FAST in China (Nan et al. 2011) and the SKA in South Africa (Keane et al. 2015) become available, the precision of pulsar timescales will improve dramatically.

Summary

Pulsars are fantastic natural clocks with a wide range of applications. Millisecond pulsars in particular have period stabilities that approach those of the best atomic frequency standards over long time intervals. MSPs in binary orbit with another star, especially another neutron star, are remarkable systems for testing theories of gravitation. In particular, the original Hulse-Taylor binary pulsar, PSR B1913+16, and the Double Pulsar, PSR J0737−3039A/B, have provided the most stringent tests, confirming that Einstein's general theory of relativity accurately describes the motion of bodies in strong gravitational fields as well as giving the first observational evidence for gravitational waves.

Pulsar timing arrays have the potential to make a direct detection of low-frequency gravitational waves. Current limits on the strength of a nanohertz gravitational-wave background from binary supermassive black holes in distant galaxies are significantly constraining current ideas about the formation and evolution of these supermassive black holes.

Pulsar timing array data sets can also be used to establish a "pulsar timescale" that can detect or limit irregularities in the best atomic timescales over long time intervals. Compared to atomic timescales, pulsar timescales:

- Are independent of terrestrial timescales and essentially independent of the terrestrial environment
- Are based on entirely different physics, namely, rotation of a macroscopic object
- Will continue without interruption for billions of years

In the future, pulsar timescales will provide a valuable "secondary standard" to the international frequency and time standards, allowing an independent check on the long-term stability of the terrestrial atomic timescales over decades and indefinitely into the future.

Acknowledgements We thank Demetrios Matsakis for providing us with the 4-year data sets for the rubidium fountain clocks at US Naval Observatory. We also thank Michael Kramer and Ryan Shannon for providing Figs. 30.2 and 30.3, respectively. The IPTA is a consortium of three PTAs, and we acknowledge the efforts of all members of these PTAs in the creation of the IPTA data sets. The Parkes radio telescope is part of the Australia Telescope which is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO.

References

B.P. Abbott, R. Abbott, T.D. Abbott, M.R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, et al., GW151226: observation of gravitational waves from a 22-solar-mass binary black hole coalescence. Phys. Rev. Lett. **116**(24), 241103 (2016)

Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. Chamberlin, S. Chatterjee, B. Christy, J.M. Cordes, N. Cornish, et al., The NANOGrav nine-year data set: Observations, arrival time measurements, and analysis of 37 millisecond pulsars. Astrophys. J. **813**, 65 (2015)

D. Bhattacharya, E.P.J. van den Heuvel, Formation and evolution of binary and millisecond radio pulsars. Phys. Rep. **203**, 1–124 (1991)
M. Burgay, N. D'Amico, A. Possenti, R.N. Manchester, A.G. Lyne, B.C. Joshi, M.A. McLaughlin, M. Kramer, J.M. Sarkissian, et al., An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. Nature **426**, 531–533 (2003)
G. Desvignes, R.N. Caballero, L. Lentati, J.P.W. Verbiest, D.J. Champion, B.W. Stappers, G.H. Janssen, P. Lazarus, et al., High-precision timing of 42 millisecond pulsars with the European pulsar timing Ar- ray. MNRAS **458**, 3341–3380 (2016)
W.M. Folkner, Planetary and Lunar Ephemeris DE418. Tech. Rep. IOM343R-07-005, NASA Jet Propulsion Laboratory (2007)
W.M. Folkner, R.A. Park, R.A. Jacobson, Planetary Ephemeris DE435. Tech. Rep. IOM 392R-16-003, NASA Jet Propulsion Laboratory (2016)
R.S. Foster, D.C. Backer, Constructing a pulsar timing array. Astrophys. J. **361**, 300 (1990)
R.W. Hellings, G.S. Downs, Upper limits on the isotropic gravitational radiation background from pulsar timing analysis. Astrophys. J. **265**, L39 (1983)
G. Hobbs, A. Archibald, Z. Arzoumanian, D. Backer, M. Bailes, N.D.R. Bhat, M. Burgay, S. Burke-Spolaor, D. Champion, I. Cognard, et al., The international pulsar timing array project: Using pulsars as a gravitational wave detector. Class. Quant. Grav **27**(8), 084013 (2010)
G. Hobbs, W. Coles, R.N. Manchester, M.J. Keith, R.M. Shannon, D. Chen, M. Bailes, N.D.R. Bhat, et al., Development of a pulsar- based timescale. MNRAS **427**, 2780–2787 (2012)
G.B. Hobbs, R.T. Edwards, R.N. Manchester, Tempo2, a new pulsar-timing package—I. An overview. MNRAS **369**, 655–672 (2006)
R.A. Hulse, J.H. Taylor, Discovery of a pulsar in a binary system. Astrophys. J. **195**, L51–L53 (1975)
E. Keane, B. Bhattacharyya, M. Kramer, B. Stappers, E.F. Keane, B. Bhattacharyya, M. Kramer, B.W. Stappers, et al., A cosmic census of radio pulsars with the SKA. Advancing Astrophysics with the Square Kilometre Array (AASKA14), 40 (2015)
B. Kocsis, A. Sesana, Gas-driven massive black hole binaries: Signatures in the nHz gravitational wave background. MNRAS **411**, 1467–1479 (2011)
M. Kramer, I.H. Stairs, R.N. Manchester, M.A. McLaughlin, A.G. Lyne, R.D. Ferdman, M. Burgay, D.R. Lorimer, et al., Tests of general relativity from timing the double pulsar. Science **314**, 97–102 (2006)
A.G. Lyne, M. Burgay, M. Kramer, A. Possenti, R.N. Manchester, F. Camilo, M.A. McLaughlin, D.R. Lorimer, et al., A double-pulsar system: A rare laboratory for relativistic gravity and plasma physics. Science **303**, 1153–1157 (2004)
R.N. Manchester, The international pulsar timing array. Class. Quant. Grav. **30**(22), 224010 (2013)
R.N. Manchester, G. Hobbs, M. Bailes, W.A. Coles, W. van Straten, M.J. Keith, R.M. Shannon, et al., The parkes pulsar timing array project. PASA **30**, e017 (2013)
R.N. Manchester, G.B. Hobbs, A. Teoh, M. Hobbs, The Australia telescope national facility pulsar catalogue. Astron. J. **129**, 1993–2006 (2005)
R. Nan, D. Li, C. Jin, Q. Wang, L. Zhu, W. Zhu, H. Zhang, Y. Yue, L. Qian, The five-hundred-meter aperture spherical radio telescope (FAST) project. Int. J. Mod. Phys. D **20**, 989–1024 (2011)
S. Peil, J.L. Hanssen, T.B. Swanson, J. Taylor, C.R. Ekstrom, Evaluation of long term performance of continuously running rubidium fountains. Metrologia **51**, 263–269 (2014)
B.B.P. Perera, M.A. McLaughlin, M. Kramer, I.H. Stairs, R.D. Ferdman, P.C.C. Freire, A. Possenti, R.P. Breton, et al., The evolution of PSR J0737−3039B and a model for relativistic spin precession. Astrophys. J. **721**, 1193–1205 (2010)
G. Petit, in Proc. EFTF-FCS'07 Meeting, Long-term stability and accuracy of TAI (revisited) (2007), pp. 391–394

E.S. Phinney, A practical theorem on gravitational wave backgrounds. ArXiv:astro-ph/0108028 (2001)
V. Ravi, J.S.B. Wyithe, R.M. Shannon, G. Hobbs, R.N. Manchester, Binary super- massive black hole environments diminish the gravitational wave signal in the pulsar timing band. MNRAS **442**, 56–68 (2014)
D.J. Reardon, G. Hobbs, W. Coles, Y. Levin, M.J. Keith, M. Bailes, N.D.R. Bhat, S. Burke-Spolaor, et al., Timing analysis for 20 millisecond pulsars in the parkes pulsar timing array. MNRAS **455**, 1751–1769 (2016)
A. Sesana, F. Shankar, M. Bernardi, R.K. Sheth, Selection bias in dynamically-measured super-massive black hole samples: consequences for pulsar timing arrays. ArXiv:1603.09348 (2016)
R.M. Shannon, J.M. Cordes, Assessing the role of spin noise in the precision timing of millisecond pulsars. Astrophys. J. **725**, 1607–1619 (2010)
R.M. Shannon, V. Ravi, L.T. Lentati, P.D. Lasky, G. Hobbs, M. Kerr, R.N. Manchester, W.A. Coles, et al., Gravitational waves from binary supermassive black holes missing in pulsar observations. Science **349**, 1522–1525 (2015)
J.H. Taylor, L.A. Fowler, P.M. McCulloch, Measurements of general relativistic effects in the binary pulsar PSR 1913+16. Nature **277**, 437 (1979)
J.P.W. Verbiest, L. Lentati, G. Hobbs, R. van Haasteren, P.B. Demorest, G.H. Janssen, J.B. Wang, G. Desvignes, et al., The international pulsar timing array: first data release. MNRAS **458**, 1267–1288 (2016)
J.M. Weisberg, D.J. Nice, J.H. Taylor, Timing measurements of the relativistic binary pulsar PSR B1913+16. Astrophys. J. **722**, 1030–1034 (2010)

Chapter 31
The Leap Second Debate: Rational Arguments vs. Unspoken Unease

Pavel Gabor

Abstract The ancient and sacred task of timekeeping, linking the eternal with the everyday, is one of the oldest missions of astronomy, originating long before the dawn of written history. The succession of timekeeping schemes throughout millennia has been a search for a balance between the practical and the symbolic. In the current Leap Second Debate, there are rational arguments, focused on practical considerations, and there is a certain unspoken unease, emerging from the symbolic substrata of the issues involved.

Developing our work presented at the Exton and Charlottesville colloquia, this paper will examine the presuppositions and perceptions overshadowing the debate: astronomical conformity, continuity, timelessness, (ir)rationality, the Enlightenment, etc. We shall study the historical evidence represented by the various calendric traditions and situate the Leap Second Debate in a broader context of cultural history and history of thought. Our underlying purpose is to facilitate the debate by shedding some light on its caliginous but potent undercurrents.

Keywords Definition of meter • History of calendars • Gregorian calendar • Enlightenment • Social alienation • Natural • Anthropic • Universal • Rational

Introduction

It is a rare privilege to share one's reflections with such an eminent audience. But it is the rarest of privileges to be able to present an update 5 years later.

The one thing that all systems of weights and measures (metric, imperial, CGS, MKS, SI) have in common is the unit of time. Time is the most mysterious of physical quantities. It is the only one that flows; it flows inexorably: We have no practical control over it. Time as a coordinate has to be handled differently from its spatial counterparts. It is a constant reminder of our mortality. And perhaps because of this we want to believe that our time, our lifetime, somehow corresponds to the eternal cosmic cycles. If our lives are intertwined with eternal cosmic cycles, we

P. Gabor (✉)
Vatican Observatory, Vatican City State, Rome
e-mail: pgabor@as.arizona.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_31

feel reassured. The cosmic cycles presumably have some unfathomable purpose and meaning; they bear a profound symbolism (although we are unaware of the meaning).

When I gave my presentation at Exton (Gabor 2011), my position was that "the principle of astronomical conformity is crucial in civil timekeeping for reasons that go far beyond the purely practical." My analysis used the history of calendars to describe several of the underlying social mechanisms that have influenced timekeeping schemes over the millennia. My conclusion regarding leap second insertion was that even if it is suspended, thus effectively decoupling civil time from Earth's rotation, such "a departure from [astronomical conformity] is unlikely to last" for more than a few centuries and that this fundamental principle of timekeeping symbolism will eventually reaffirm itself.

Today I would like to add a few remarks about some of the social forces that have shaped our attitudes toward metrology and standards. Again, I shall advocate the view that the big picture matters. Facing the daunting and intriguing challenge of addressing a multidisciplinary audience, I decided to start with a story. I believe that a narrative may be the best vehicle for my message. My story may be inaccurate in some details, but the purpose is to convey a message, a view of the big picture, which I believe to be true and pertinent.

The Brave New World

Once upon a time, the world was violent and turbulent. There were wars and turmoil under the pretext of religious affiliation. The worst of these wars lasted 30 years. Battlefield casualties were not particularly high but devastation due to plague and famine was unimaginable. The so-called total wars of the twentieth century killed a much smaller proportion of Europeans than these sectarian wars did.

After a couple of centuries of mayhem, the Enlightened ones said, "Let us build a brave new world upon the foundation of Reason—the one commonality of all human beings." They felt this to be natural. Does not the philosopher say that man is *animal rationale*? After all, the Enlightened thinkers of the eighteenth century all came from 500-year-old tradition of Aristotelian schools which taught that reason was the key feature of human beings, distinguishing them from other animals.

The "religious wars" were so horrific that the Enlightened reforms became attractive to a critical mass of thinkers, politicians, princes, and other potentates.

The Measure of All Things

An important (if not particularly glamorous) project within this effort of creating a new and better world was the reform of weights and measures. In this mindset, the old was bad simply because it had been associated with the old world order. Age was a sufficient reason for rejection. The new world needed new units.

Furthermore, in the old world, each place, each political entity had its own weights and measures. In the eyes of the architects of the Enlightened world, this was yet another symptom of the ills of the old order: The units of measure perpetuated divisions among people, fostering the sectarian mentality. In the new age of Reason, everybody was to accept the one system that was the most reasonable of all imaginable systems.

Finally, most of the old units were anthropic: the inch, the foot, the fathom, etc. The new units were to be universal and derived from nature herself.

The grand plan, therefore, was to reject the old, local, and anthropic standards and introduce new ones perceived as objective, natural, universal, and rational.

It may be useful to mention that the relationship between truth and reality on the one hand and their *perception* by individuals and societies is curious. By and large, humans *define* their "reality": What is *perceived* as real and true is more important than what really is real and true in an objective sense. Physical reality provides some boundaries but, surprisingly, it can be very often ignored with impunity.

The shift from the particular and anthropic units to "universal" and "natural"[1] units is part of the symbolic drive toward the pretense of rationality and objectivity. Man was no longer to be *the measure of all things* (as Protagoras's logion goes[2]).

Ever since then, we have been gradually moving away from the human-friendly to the "scientific" (to use current language). The latter may be better as judged by some pertinent metrics, but it causes social alienation.[3] Although this is (for the most part) quite unintentional, science powerfully meddles with the collective unconscious.

Let us return to metrological standards. The community of metrologists institutionally values standardization. After all, the universal adoption of the same units has been the worthy goal in this field at least since the eighteenth century (we shall see some illustrations in the following section). And metrological consistency is a

[1] I put this pair of adjectives in quotation marks for two reasons. (1) These terms were actually used at the time, as we shall see in the following section. (2) These attributes are misplaced: The new system was hardly universal or natural in actual fact.

[2] Protagoras of Abdera probably meant practically the opposite of what people using the logion assume. The philosopher's point was that humans define their own truth and reality: What counts is how I perceive things.

[3] Here is a definition of this sociological term: Alienation is "a condition in social relationships reflected by a low degree of integration or common values and a high degree of distance or isolation between individuals, or between an individual and a group of people in a community or work environment" (Ankony 1999).

very attractive goal even today. I could put it no better than ITU's Working Party 7A:

> UTC without leap seconds will represent a continuous reference timescale and will encourage the use of only one continuous reference timescale making it truly universal. This will avoid the proliferation of other timescales that may cause serious confusion and contribute to the interoperability of systems. (Working Party 7A 2015)

The presupposition underlying this paragraph is that people must adopt a single timescale because there would be "serious confusion" if they did not. Let me point out that the imposition of uniformity is a symbolic demonstration of power. It belongs to an entirely different area of symbols than timekeeping does. In order for a timekeeping scheme to maintain its symbolic power, it does not require to have a monopoly in a given society. History teaches us that several timekeeping systems may coexist within the same community without challenging each other. Imposing uniformity is a symbolic gesture linked to power, conquest, and domination. And paradoxically, it is diametrically opposed to fostering genuine organic unity.

The Meter

The story of the meter is an excellent example of such an undertaking. For a while, there was the elegant idea to adopt the length of the one-second pendulum as the base unit. It was officially proposed to the French Assembly on 9 March 1790 by Charles Maurice de Talleyrand-Périgord, then Bishop of Autun.[4] On 8 May 1790, the Assembly adopted the principle. On 26 March 1791, however, the Marquis Nicolas de Condorcet read to the Assembly a report prepared by the Academy of Sciences, stating:

> The Academy has sought to exclude any arbitrary conditions, anything that could elicit the suspicion of particular French national interests or prejudice. To put it simply, she wishes that if somehow only the principles and specifications of this operation pass to posterity, it would be impossible to tell which nation had carried it out. [. . .] It is important to choose a system which could suit all peoples. The success of the undertaking depends to such a point on how general its foundations will be that the Academy does not judge that it could merely refer to measurements already performed, nor can she be satisfied with simply observing the pendulum. The Academy feels that, working for a powerful nation, on the orders of enlightened men who embrace in their sight all men and all the ages, she must attend less to what would be easier than to what would be closest to perfection. The Academy believes that she would not be unworthy to accept such a great undertaking manifesting the enlightened zeal of the National Assembly for the growth of light, progress and the fraternity of nations. (Février 2014, my translation)

This offer was accepted, and on the same day, the Assembly passed a decree (approved by the King on 30 March) stipulating:

[4]Talleyrand was later foreign minister under Napoleon, Louis XVIII, and even Louis-Philippe.

> Whereas, in order to arrive at establishing the uniformity of weights and measures, it is necessary to set a natural and invariable unit of measure, and whereas the only means of extending this uniformity to foreign countries and of committing them to a new measuring system is to choose a unit which encompasses nothing arbitrary nor particular to the situation of a single nation of the world [...] [the Assembly] adopts the length of the Earth's meridian quadrant as the base of the new system of measurements which shall be decimal. The work required to determine this base, namely the measurement of a meridian arc between Dunkirk and Barcelona, will be executed without delay. (Février 2014, my translation)

I find the rhetoric fascinating. The presumption of these people! They had less knowledge of the physical universe than the average freshman today (and that's saying something!), but they had the confidence of a 4-year-old who thinks he's just figured out how the world works. They actually believed themselves capable of devising a "natural and invariable unit of measure" with "nothing arbitrary" about it, a wholly new system, defined once and for all, for "all men and all the ages."

By the way, notice how the Academy subtly rubbishes the 1-second pendulum idea. It was not glamorous enough for such a "great undertaking" as the determination of this ultimate unit[5] simply had to be: It had to involve expending huge amounts of time and effort—whether they were necessary or not. Some say, however, that the decisive argument had to do with the second itself: It was not a decimal unit. Why should it be in the definition of the base unit of length in a new decimal system?

In any case, it is morbidly fascinating to observe how these people, most of whom would literally worship Reason (Fig. 31.1), were so taken in by their own rhetoric. Today, it is hard to see how anyone could have taken the metric enterprise seriously and ignored its deep methodological flaws. Of course, with the benefit of hindsight, we know that it soon became clear that the meter was, in truth, an arbitrary unit. It does not really correspond to the ten-millionth part of the quarter of Earth's meridian—even if that definition could be, by some stretch, considered "universal" and "natural."

The Gregorian Reform

I could give many more arguments why I think that the unbridled rationalist and progressist zeal of the Enlightenment is not a legacy with which we should identify. Human society cannot stay excited about anything for more than a generation. It saddens me to say this, but my observations of ordinary people lead me to believe that they do not perceive the meter as a "natural" unit, but rather as one example among many arbitrary impositions made by faceless experts who "know what is

[5]This can be felt in the choice of the verb, "ni se contenter de la simple observation du pendule." My English translation, "nor can she be satisfied with simply observing the pendulum," does not quite convey the nuance.

Fig. 31.1 Translation of the inscription under the engraving: "On the tenth day [of the ten-day 'week'], 20th of the Misty month of the second year of the French Republic one and indivisible [10 November 1793], the Feast of Reason was celebrated in the erstwhile Church of Our Lady [in Paris]." The engraving shows a symbolic mountain with a Greek temple dedicated to philosophy. At its base there was an altar dedicated to Reason. The torch of truth is in the foreground. The ceremony included homage to an actress dressed in blue, white, and red (the colors of the Republic), representing personified Liberty. (Wikimedia)

best for us." Today, 220 years after the events described in the previous section, the social alienation perceived as "caused by science" is not diminishing. Far from it. It is worse today than it was just a few decades ago.

I would like to point out that in early modern times, there was a major and dramatic reform of timekeeping that has caused much upheaval, protest, and division: the Gregorian reform of the calendar (Fig. 31.2). More than four centuries after *Inter gravissimas*,[6] we can safely say that society perceives the calendar as conforming to astronomical cycles.

Is this perception accurate? As I have pointed out in Exton, "Symbols work as long as they are *perceived* as grounded in reality: Timekeeping symbols work as

[6]The reform's Bull of promulgation, issued by Gregory XIII on 24 February 1582. Note that this date was carefully chosen: In Latin it is *ante diem sextum kalendas Martias*. March is the first month of the year in the Roman calendar (September is the seventh, October the eighth, November the ninth, and December the tenth; January is the eleventh month, and the last month is February). The leap day (introduced by Julius Caesar) is inserted between the sixth and fifth day before the calends of March (1 March), and hence it is *ante diem bis sextum Kalendas Martias* (duplicated sixth day before 1 March).

Fig. 31.2 The tomb of Gregory XIII in St. Peter's Basilica, Vatican. The committee for the reform of the calendar presents its report to the Pope. (Wikimedia)

long as they conform to astronomical phenomena at least in some way that would allow the general perception to persist." In the case of the Gregorian calendar, the intention of the legislator is very clear: He intends to ensure that the calendar conform to astronomy (e.g., *cavetur, ut in posterum quinoctium et XIV luna a propriis sedibus numquam dimoveantur*[7]). It is this declaration of intent that carries the day, regardless of any factual imperfection or discrepancy.

Let us examine the matter in detail. The Bull introduces one mechanism *ne in posterum a XII Kalendas Aprilis æquinoctium recedat* (dropping 3 leap days in every 400 years), but it does not mention the fact that an additional tweak will be needed in a few thousand years' time. Thus, if we were to regard *Inter Gravissimas* as the complete definition of the Gregorian reform, then the new calendar would be inherently astronomically nonconformant, however small may its inaccuracy be. But the Bull was never intended to provide a complete description of all the details.

The practice of the Roman Curia is quite consistent with this. A Bull is a concise text and it is not expected to contain everything. It is a statement of principle, and the implementation and other fine points are rarely included. A Bull *marks* a decision: It is not meant to describe everything in exhaustive detail. A Bull is not the decision: It announces that a decision has been made. The legislator is a single person, and the decision is not a legislative process with bills and acts of parliament, signing into law, etc. The decision simply coincides with the will of the Supreme

[7] . . . so that the equinox does not recede from 21 March in the future . . .

Pontiff. A Bull only documents the decision: The decision always precedes the document. And many decisions, even today, are undocumented.

Therefore, if we want to know how exactly is the calendar supposed to work, we need another source. It so happens that there is such an explanatory text for us to consult. Its title is *Romani Calendarii a Gregorio XIII Pontifice Maximo Restituti Explicatio*, i.e., "Explanation of the Roman Calendar restored by the Supreme Pontiff Gregory XIII" (Clavius 1603). In it Christoph Clavius of the Society of Jesus describes and discusses all possible aspects of the calendar in minute detail. This tome of some 600 pages was perceived as the quasi-official report of Gregory XIII's committee. The *Explicatio* states that several tweaks will be needed in the future adjusting the calendar to the length of the tropical year, to the precession of equinoxes, to the motion of the moon (table of epacts), etc. Clavius very openly admits that many parameters are not known with sufficient uncertainty and that the calendar cannot be defined more accurately until more and better data become available.

Clavius acknowledges the limitations of his knowledge and leaves future generations to perfect the calendar once a tweak is needed. It is an admission of the obvious, but by no means, it is an admission obvious to make.

Conclusion

Since 1999 we have witnessed a most fascinating discussion about the leap second, and we know that it will not be over until at least 2023. Many of you have been involved in the debate for years. You may have set out assuming this was purely a technical issue. And some of you may now suspect that the problem has other dimensions: There is too much unspoken unease for it to be otherwise.

My conclusion today will address those of you who are in favor of the McCarthy and Klepczynski (1999, pp. 50–57) proposal. I am not in favor of it myself, but let me suggest a few ways how you might avoid unintentionally creating more unspoken unease. You have not convinced the general public, and you may never succeed unless you avoid pushing the society's subliminal buttons.

Suspending the leap second mechanism might be the most expedient measure. But I urge you to realize that this is not just a technical matter. Decoupling civil timekeeping from Earth's rotation is a sensitive issue. How can it be done minimizing its impact on society? Never imply that the only avenue open to rational minds is to agree with the technical arguments. There are many nontechnical dimensions of our lives and they can all be discussed rationally.

The key is public perception. It is imperative to allow the general public to perceive civil time as a reflection of astronomical phenomena. In practice, claim that you intend to maintain the coupling between civil time and Earth's rotation in the long run. Never speak of "abolishing" or "suppressing" or "discontinuing" the leap second. Always maintain that the measure is temporary. Nobody should

presume to create an "ultimate" solution. The message should say, "inserting leap seconds will be suspended until a better solution is found."

Let me repeat that I am not in favor of the proposal myself. I am just trying to reduce the unnecessary unspoken unease in the hope of facilitating the discussion.

References

R.C. Ankony, The impact of perceived alienation on police officers' sense of mastery and subsequent motivation for proactive enforcement. Policing: Int. J. Police Strateg. Manage. **22**(2), 120–132 (1999)

C. Clavius, *Romani Calendarii a Gregorio XIII. P. M. Restituti Explicatio. Clementis VIII. P. M. iussu edita* (Aloysius Zannettus, Roma, 1603)

D. Février, *Histoire du mètre*, 2014. http://www.entreprises.gouv.fr/metrologie/histoire-metre

P. Gabor, in *The Heavens and Timekeeping, Symbolism and Expediency*, ed. by J.H. Seago, R.L. Seaman, S.L. Allen. Decoupling Civil Timekeeping from Earth Rotation (American Astronautical Society, 2011)

D. McCarthy, W. Klepczynski, GPS and Leap Seconds. Time to Change? *GPS World*, 1999, pp. 50–57

Working Party 7A. Agenda Item 1.14. *2nd Session of the Conference Preparatory Meeting for World Radiocommunications Conference 2015*, 2015, p. 57

Chapter 32
How to Talk to the Public About the Leap Second? The Experience of the IERS Central Bureau

Wolfgang R. Dick

Abstract The increased interest of the public in UTC and the leap second demands popular explanations on its background, practice, and future. This paper gives some hints on how to talk about this matter to different audiences and to journalists. It concentrates on the following topics:

1. Background: Explaining the difference between the slow increase of the length of day (LOD) with time on the one hand and the rather large frequency of leap seconds on the other hand
2. Prospects: The future of the time system with and without leap seconds
3. History: The end of the Earth as the most precise clock around 1935

Keywords Leap second • Public relations • Length of day • History of time measurement • Quartz clocks

Introduction

During recent years, and especially in 2015, the interest of the public in the leap second has increased considerably. This was due to the discussions about the redefinition of Coordinated Universal Time (UTC) and claims that the leap second may threaten computer systems. Also the number of users of leap second information (Bulletin C and other formats) has increased over the last years, reflecting a wider need and interest in precise timekeeping. For a discussion of these and other reasons, see Dick (2011).

An article about the possible threat by the forthcoming leap second, written by a science journalist for a popular German political journal (Dambeck 2015) seemed to be the initial source of other German journalists' interest in this matter. After the publication of this article and until June 30 (the date of the 2015 leap second), the

W.R. Dick (✉)
IERS Central Bureau, Bundesamt für Kartographie und Geodäsie, Richard-Strauss-Allee 11, 60598 Frankfurt am Main, Germany
e-mail: dick@iers.org

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_32

Central Bureau of the International Earth Rotation and Reference Systems Service (IERS) and the Federal Agency for Cartography and Geodesy (Bundesamt für Kartographie und Geodäsie, BKG) as its host organization got a large number of requests from journalists about the leap second (including the largest German news agency, from which many newspapers obtained and published two different articles), numerous invitations to interviews for radio and TV stations, individual questions from IERS users, as well as requests for public talks. There were also a large number of discussions with colleagues about the background of the leap second and its future. Colleagues at other institutions connected with time in France, Germany, and the USA have reported to us about a similar large public interest.

Following a more general discussion of the IERS, the leap second, and the public (Dick 2011), this paper will concentrate on some special topics in explaining the leap second to the public.

The Leap Second

Leap seconds are introduced as time steps into UTC, which is the basis for civil time: UTC = TAI – leap seconds (TAI = International Atomic Time), so that the last minute of a certain day has 61 s. This may occur at the end of each month, with the first preference on June 30 and December 31 (and only these two dates have been used so far). The main condition is that $|\,\mathrm{UT1–UTC}\,| < 0.9$ s (UT1 = Universal Time, the rotation angle about the Earth pole). For more details see ITU-R (2002).

This technical definition cannot be used when talking to the public; therefore, more popular explanations are needed.

The announcement of leap seconds is the responsibility of the IERS. This task was given to the IERS Earth Orientation Center at Paris Observatory, which extrapolates UT1 data into the future to decide about the next leap second and announces it in IERS Bulletin C, 5–6 months in advance. Users may subscribe for a distribution list of IERS Bulletin C.

The prediction of UT1 demands knowledge of the variable UT1 values for the past, which are obtained by observations of radio signals from cores of distant galaxies with several radio telescopes around the world. The technology for this is called very long baseline interferometry (VLBI).

The IERS Central Bureau and Leap Seconds

The work of the IERS (Dick and Richter 2004) is coordinated by a Central Bureau. It is also responsible for the collection, dissemination, and long-term storage of all IERS products, among them Bulletin C (leap second announcements), and the

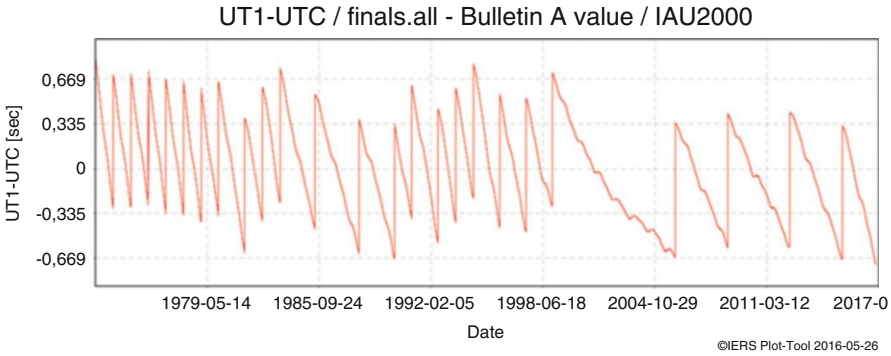


Fig. 32.1 Graph of UT1–UTC since 1973, drawn from IERS product finals.all (IAU2000) and available at https://www.iers.org/EOP. The latest jump stems from the 2015 leap second. The last part of the curve displays predicted values and shows the necessity of the next leap second, which will be introduced at the end of 2016

creation of metadata for the products. It partially manages subscriptions to Bulletin C in cooperation with the Earth Orientation Center.

The Central Bureau maintains the main IERS website www.iers.org and is responsible for contacts with users and for public relations of the IERS. It offers also several online tools, including the IERS plot tool for drawing graphs from IERS data, and it publishes at the website automatically generated graphs for IERS products. These can be used in explaining the leap second for a more technically oriented audience (Fig. 32.1).

The BKG and Leap Seconds

Since 2001, the IERS Central Bureau is hosted by the Federal Agency for Cartography and Geodesy (BKG) at Frankfurt am Main, Germany. Requests about the leap second by the public or by journalists addressed to BKG are therefore passed on to the Central Bureau, and often no clear distinction is possible whether a request is answered on behalf of IERS or of BKG.

BKG plays a major role in UT1 determination, which is the basis for the introduction of leap seconds into UTC. It operates three geodetic observatories, where—among others—VLBI observations are carried out: in Wettzell (Germany, Fig. 32.2), O'Higgins (Antarctica), and La Plata (Argentina, formerly at Concepción, Chile). Apart from the VLBI observations, BKG contributes to the International VLBI Service for Geodesy and Astrometry (IVS) by maintaining one of the IVS Data Centers, one of the Analysis Centers, as well as the IVS Combination Center, which combines all results provided by the Analysis Centers into the final IVS solution, including UT1 data.

Fig. 32.2 The three VLBI telescopes at BKG's Wettzell Observatory, Bavarian Forest, Germany. © Uwe Hessel/BKG, 2015

The author has also participated on behalf of BKG in discussions regarding the official German opinion about the new definition of UTC organized by the Bundesnetzagentur (Federal Network Agency), which represents Germany in the International Telecommunication Union (ITU), responsible for UTC definition.

Background: Explaining the Leap Second

Increase of the Length of Day

The second was originally defined for an exact duration of the day of 24 h. However, tidal friction in the Earth–Moon system increases the duration of the day. A measure for this is the so-called length of day (LOD) = duration of day—24 h. The increase of LOD is about 1.75 ms/century and currently LOD is about 1–2 ms.

A lot of physics (conservation of angular momentum, etc.) as well as astronomy, geodesy, and geophysics can be demonstrated for tidal friction when there is time and the audience is more technically and scientifically oriented. However, in most cases, it is sufficient to simply state that the Moon is slowing down the Earth through tides, and usually this is taken as granted without many questions. From our experience, it is much more difficult to explain how we come to a second from these small values of the order of milliseconds.

How We Get 1 s from 2 ms/Century?

Even astronomers and geodesists are often confused about the difference between the slow increase of the length of day (LOD) with time on the one hand and the rather large frequency of leap seconds on the other hand. To explain this difference in a popular manner, we use two clocks as a model, where one is showing the exact time and the other is too slow (Fig. 32.3). Even when the incorrect clock does not

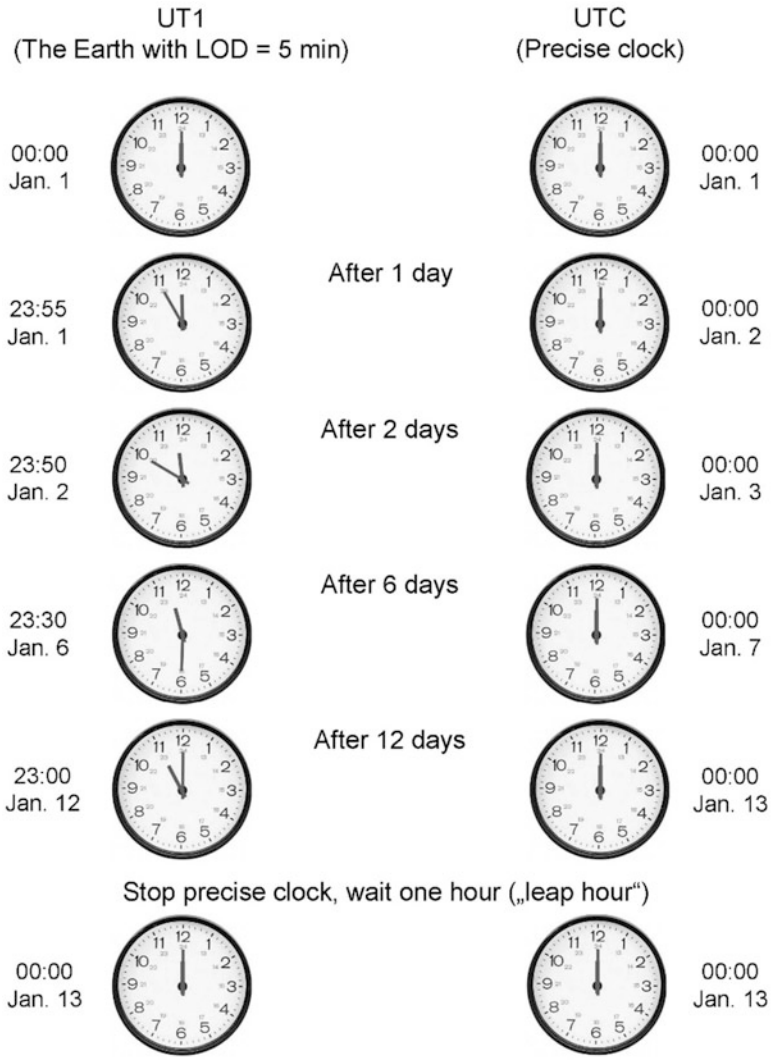


Fig. 32.3 A model for explaining the leap second with two clocks for UT1 and UTC. With an LOD of 5 min, a "leap hour" is necessary after 12 days. With an LOD of 1 ms, a leap second has to be introduced after 1000 days

change its speed, the time difference between the two clocks increases with each day. When the difference between the two times displayed will reach a certain amount, the exact clock is stopped for this amount of time to show again the same time as the slow clock. This model can then easily be transferred to UTC (exact clock) and UT1 (Earth as clock with LOD > 0).

In Fig. 32.3, LOD is set to 5 min. After 12 days the difference is already 1 h. Normally, one would put the slow clock ahead to the exact time, but this cannot be done with the Earth, so the precise clock is stopped for 1 h to display the same time as the "Earth clock."

This way it can easily be realized why an LOD of 1 ms demands a leap second after 1000 days. This can also be explained without a picture like in Fig. 32.3, because people can easily imagine two clocks, one of which is precise and the other is too slow.

Prospects: The Future of the Leap Second

Another frequently asked question is the future of the time system with and without leap seconds, especially in connection with the discussions about the proposed redefinition of UTC, about which many science journalists have read. Table 32.1 shows the mean length of the day, ΔT^*, and the mean time between two leap seconds over the next 2000 years as a very rough estimate. The mean LOD is 0 for the year 1900 and increases linearly by about 1.8 ms per century (neglecting decadal changes by core-mantle coupling, etc.). ΔT^* is the difference between TAI und UT1 (or mean solar time), i.e., between the time scale of a precise (atomic) clock and a time scale defined by the Earth as a clock which rotates slower and slower. It is very roughly estimated as

$$\Delta T^* = 32 \text{ s } T^2 \tag{32.1}$$

with T in centuries from 1900. This formula is much simplified but for a more technically oriented audience clearly shows the quadratic dependence on time. (See McCarthy and Seidelmann 2009, p. 89, and the literature cited therein for more exact considerations and formulas. Scientific studies express long-term changes in Earth rotation as ΔT, the difference between Terrestrial Time (TT) und UT1. However, for popular purposes it would be confusing for the audience to hear about another, rather abstract time scale. TT and TAI are only different by a constant.)

It can easily be seen that the current system with a maximum of 12 leap seconds per year will fail about 1500 years from now. In the year 4000, ΔT^* will be about 4 h, which means that at mid-latitudes the Sun would rise then around 12:00 in winter time if leap seconds would be abolished, the current time zones would remain unchanged, or no other measures would be taken to bring civil time in

Table 32.1 Rough estimates of mean LOD, ΔT^*, and mean time between leap seconds over the next 2000 years

Year—2000	Mean LOD	ΔT^*	Time between leap seconds
0	2 ms	32 s	500 days
100	4 ms	2 min	250 days
500	11 ms	20 min	90 days
1000	20 ms	1 h	50 days
1500	29 ms	2.3 h	35 days
2000	38 ms	4 h	26 days

The numbers in the first line are mean values over several decades around the year 2000. Due to a core-mantle coupling, the Earth speeded up around 2000 with a mean LOD of about 1 ms during the following decade, so that leap seconds occurred less often than before (cf. Fig. 32.1)

agreement with solar time. In this case the English word "noon" or the French "midi," associated with the position of the Sun in the sky, would lose their meaning.

With this number (4 h in 2000 years) it can easily be seen why it is necessary to bring civil time in agreement with solar time. It is also possible to explain the intentions behind leap seconds in a very short and very popular way: Earth is rotating slower and slower, and without the leap seconds, the Sun would rise later and later, finally only at 12 o'clock after 2000 years.

This very short statement is not quite true, because there are other possibilities to change the civil time scale to accommodate for a slowing down Earth, e.g., through a change of time zones. But this sentence seems to be a way to give an explanation of a forthcoming leap second when you are interviewed by a radio or TV journalist and have practically only seconds (one to three sentences) to answer. The original intentions for defining UTC with leap seconds (celestial navigation, etc.) cannot be explained shortly, and even when you start to explain these, it is not possible to explain easily why this system is still kept although the original reasons for introducing it have lost their importance. One has to be very careful when discussing the pros and cons of leap seconds with journalists, because they easily can get the false impression of the "typical mad scientists"—they do things (introducing leap seconds) which are no longer necessary. The easiest way to describe the current discussions among specialists about the leap second is to say that an international agreement about the best method to bring civil time in agreement with the Earth slowing down has not yet be reached.

History: The End of the Earth as the Most Precise Clock

Until the beginning of the 1930s the Earth was the most precise clock we had. Although from observations of the motion of the Moon it became clear that Earth's rotation rate must vary (McCarthy and Seidelmann 2009, pp. 44–51), it was not possible to measure this directly with the help of pendulum clocks.

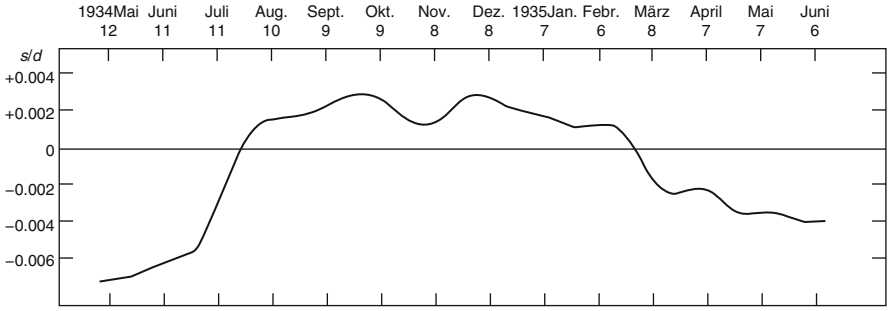


Fig. 32.4 The first published LOD graph (Pavel and Uhink 1935, Fig. 11). The order of magnitude of LOD variation and the trend are both correct

When the physicists Adolf Scheibe and Udo Adelsberger at Physikalisch-Technische Reichsanstalt (PTR) in Berlin, the forerunner of today's Physikalisch-Technische Bundesanstalt (PTB), built the first quartz clocks with constant behavior (not the first quartz clocks in general), it became possible to compare the time from these clocks with astronomically derived time. Scheibe and Adelsberger used time signals for this; they did not do their own astronomical observations.

Before 1935, two similar quartz clocks were built at the Geodetic Institute at Potsdam (close to Berlin) from plans by Scheibe and Adelsberger, and the astronomers/geodesists Friedrich Pavel and Werner Uhink also compared time from their quartz clocks with mean solar time from time signals. Pavel and Uhink were the first who published an LOD graph (see Fig. 32.4) in the journal *Astronomische Nachrichten*, hidden in a very technical paper about quartz clocks (Pavel and Uhink 1935).

In German literature, credit is given only to Scheibe and Adelsberger; e.g., the German Wikipedia page https://de.wikipedia.org/wiki/Adolf_Scheibe names Scheibe the "discoverer of the irregularity of Earth rotation", the page https://de.wikipedia.org/wiki/Udo_Adelsberger even claims that Scheibe und Adelsberger "shocked the specialists" with their discovery. These statements are taken in modified form from printed sources on the history of the PTB and on Scheibe and Adelsberger, in which Pavel and Uhink are neglected (as did Scheibe and Adelsberger in their publication of 1936). That the specialists were "shocked" is not true. Pavel and Uhink commented their result with a statement that this may be the variability of Earth rotation suspected since long ago ("... daß hier die schon lange vermutete Ungleichmäßigkeit in der Rotation der Erde gefunden worden ist," Pavel and Uhink 1935, column 389; see also Bauch 2012 for a better historical review of the work at PTR/PTB). But it must have been a shock to Scheibe and Adelsberger that another group published the first positive result before them, using technology developed by them. They submitted a short notice with their own results in December 1935 and published a longer paper the next year (Scheibe and Adelsberger 1936a, b). It would be worth investigating more thoroughly the

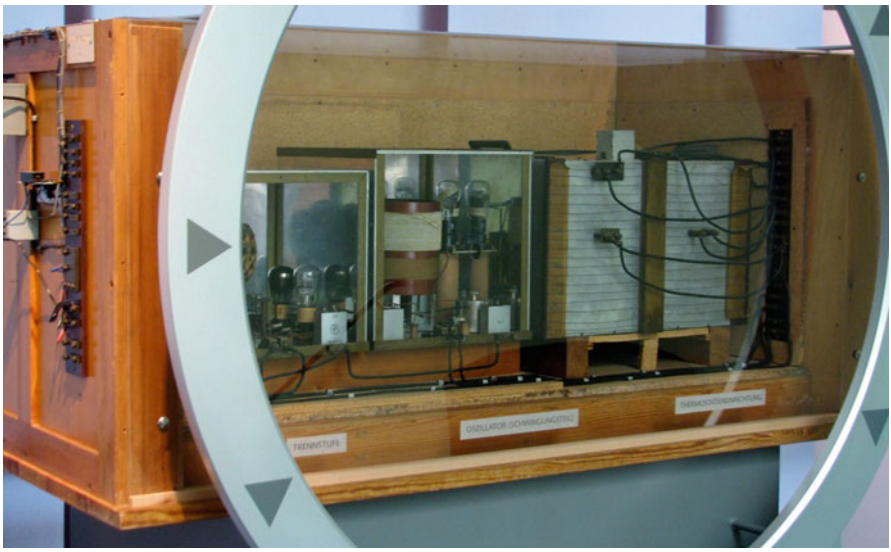


Fig. 32.5 The remaining part of one of the quartz clocks of Potsdam Geodetic Institute, built in the 1930s, on display in the Städtisches Museum Zeulenroda, Thuringia, Germany. © The author, 2014

relations of the two groups and the details of the first direct observational proofs of variability of Earth's rotation.

Unfortunately, the original quartz clocks of the PTR are lost. Half of a similar quartz clock built at the Potsdam Geodetic Institute in the 1930s has been preserved and is now on display in the city museum of Zeulenroda, Thuringia, Germany, the hometown of A. Scheibe (Fig. 32.5).

It took about 25 years before atomic time replaced time scales derived from Earth's rotation for civil timekeeping. However, the first direct observational proof for the variability of Earth's rotation was found in 1935 in Potsdam and Berlin by two groups independently from each other.

Acknowledgments For discussions and a critical reading of the first versions, I would like to thank Dr. John Bangert, Dr. Galina Dick, Catherine Hohenkerk, and Dr. Daniela Thaller; for providing the cited publication on the history of time measurement at PTB: Dr. Andreas Bauch; for information on the quartz clock at Zeulenroda museum: Dr. Johannes Leicht; for technical and other help with figures: Sonja Geist, Dr. Anja Niederhöfer, and Sandra Schneider-Leck.

References

A. Bauch, Zeitmessung in der PTB. PTB-Mitteilungen **122**(1), 23–36 (2012)

H. Dambeck, Umstellung im Juni: Schaltsekunde bedroht Computersysteme (2015). Spiegel Online. http://www.spiegel.de/wissenschaft/weltall/weltzeit-schaltsekunde-bedroht-weltweit-it-systeme-a-1011914.html. Accessed 22 July 2016

W.R. Dick, The IERS, the Leap Second, and the Public, in *Decoupling of Civil Timekeeping from Earth Rotation. Proceedings of a Colloquium Exploring Implications of Redefining Coordinated Universal Time (UTC), held October 5–7, 2011 at Analytical Graphics, Inc., Exton, Pennsylvania*, ed. by J.H. Seago, R.L. Seaman, S.L. Allen. American Astronautical Society, Science and Technology Series, vol. 113 (Univelt, San Diego, CA, 2011), pp. 117–122

W.R. Dick, B. Richter, The International Earth Rotation and Reference Systems Service (IERS), in *Organizations and Strategies in Astronomy*, vol. 5, ed. by A. Heck. Astrophysics and Space Science Library, vol. 310 (Kluwer Academic Publishers, Dordrecht, Boston, London, 2004), pp. 159–168

ITU-R, Recommendation ITU-R TF.460-6: Standard-Frequency and Time-Signal Emissions (2002). https://www.itu.int/rec/R-REC-TF.460-6-200202-I/en. Accessed 4 Aug 2016

D.D. McCarthy, P.K. Seidelmann, *Time – From Earth Rotation to Atomic Physics* (Wiley-VCH Verlag, Weinheim, 2009)

F. Pavel, W. Uhink, Die Quarzuhren des Geodätischen Institutes in Potsdam. Astronomische Nachrichten **257**, 365–390 (1935)

A. Scheibe, U. Adelsberger, Nachweis von Schwankungen der astronomischen Tageslänge mittels Quarzuhren. Physikalische Zeitschrift **37**, 38 (1936a)

A. Scheibe, U. Adelsberger, Schwankungen der astronomischen Tageslänge und der astronomischen Zeitbestimmung nach den Quarzuhren der Physikalisch-Technischen Reichsanstalt. Meßergebnisse II. Physikalische Zeitschrift **37**, 185–203 (1936b)

Chapter 33
The Problem of Leap Seconds

Bob Frankston

Abstract A leap second is an adjustment made periodically to keep UTC (Coordinated Universal Time) aligned with rotation of the Earth because the length of a day isn't exactly 86,400 s. Without such an adjustment, in 5000 years, noon will be an hour off from the actual position of the sun in the sky. As I'll explain, a better approach would be a once every 5000 years adjustment akin to the annual daylight adjustments to the clock.

As I'll explain, the current practice of inserting seconds, in effect, redefines the minute. A minute whose length is uncertain causes unnecessary collateral damage. Applications that need the precision depend on finger-grained adjustments (such as UT1). Those applications cannot depend on civil time because, in practice, one cannot determine if a time stamp actually takes leap seconds into account. The best approach is to eliminate leap seconds so that those applications that don't require the precision aren't confused and those that do won't be at risk due to the uncertainties in civil times.

Keywords Coordinated Universal Time • Leap second • Civil time

Second Thoughts

From The Week magazine:

> **Good week for**... **Everyone who needs more time,** after NASA [sic] announced that a "leap second" will be added to the world's official clocks at 11:59:59 on New Year's Eve this year, an adjustment meant to keep atomic clocks in sync with the Earth's rotation.

(Note that the leap second is decided by the International Earth Rotation and Reference Systems Service not NASA.)

B. Frankston (✉)
Independent scientist, Boston, USA
e-mail: ScienceOfTime2016@bob.ma; ScienceOfTime@bob.ma; http://Frankston.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_33

What does this announcement mean? An atomic clock doesn't have the concept of minutes or the concept of the Earth's rotation. It's just a counter that runs at a constant rate independent of the rotation of the Earth.

Clocks are used to synchronize events. The wall clock showing 11:59:59 (PM) is used to synchronize people's activities. Synchronizing with the position of the sun is a separate requirement. Where the precise position matters, we don't use the wall clock. Instead we rely on the atomic clock which doesn't have the concept of a minute. In 5000 years, people may care that noon is more than a time zone (hour) off at which point we can do a one-time renaming (as with daylight savings time). Until then making the hand on the clock jump to accommodate the leap second just adds confusion and creates unnecessary problems.

(See https://en.wikipedia.org/wiki/Leap_second for examples of the problems.)

What Is a Leap Second?

The term is borrowed from "leap year." Every 4 years (with some exceptions) there is a 29th day added to February to compensate for the fact the year is actually 365¼ days long. This means, for example, that you can't calculate the number of days between February 28 and March 1 unless you know what year it is.

A leap second is added (or, theoretically, could be subtracted) to compensate for the changes in the Earth's rate of rotation. That second has to be put somewhere so it is assigned to an arbitrary minute (usually the last one in June or December). This has consequences. A minute is no longer defined as 60 s! That extra second that is added or removed counts toward the length of that 59th minute. You can no longer say that there are 3600 s in an hour! All school exams that involve such calculations are wrong!

You can't redefine the minute any more than you can pass a law changing the value of π. As we'll see there is no need to change the definition of the minute because calculations in computers are done in seconds (or fractions of a second). The command to add 1 min always adds 60 s. It doesn't know which minute so the standard libraries can't even represent a minute with 61 s (Fig. 33.1).

Longer term, we can compensate for the slowing of the Earth's rotation by adding an hour once every 5000 years or so.

Consequences

The leap second comes at a price. It puts a lie to the very standard we depend on in the modern world. NIST defines the minute as 60 s. But that's wrong thanks to leap seconds. If the standards organization is confused, imagine how confused the rest of us are!

Fig. 33.1 Clock showing the 61st second of a minute

Banks typically shut down their computers during the leap-second transition because they can't risk what will happen when their clocks or, more to the point, only some of their clocks don't jump and become out of sync.

There is also the lurking bug. When a computer stores the year, it represents the value as seconds since from a standard base time (often called an epoch). The value computed for the year 2030 now (in 2016) will not be the same as the value computed in 2030 because we don't know how many leap seconds will be added in that interval. When a computer compares two values it will find that two values for 2030 are different! In practice this doesn't happen because most databases programs have a bug and simply ignore the leap second. But if that bug ever gets fixed we'll discover that we've got a Y2K like bug in the heart of every database system!

Do we really want to depend on a bug hiding a bug in order to make this all work?

And, as we'll see, all this complexity and uncertainty is simply so we can align the wall clock to the position of the sun. It is an idea whose time has passed. The clock is now coupled to human events and not the heavens.

Defining the Second

A cesium atomic clock counts the hyperfine structure transitions of the cesium-133 atom. A second is defined as 9,192,631,770 transitions so, in effect, the clock counts seconds. Computers also count seconds. As I write this, it is 1,467,487,440 s since January 1, 1970 (the Unix epoch or starting date). International Atomic Time or Temps Atomique International (TAI) operates with Julian days. This is not the same as the Julian calendar but is simply a count of days starting at an arbitrary date. January 1, 2013 is Julian day 2,456,293.

Prior to the use of the atomic clock, a second was defined as 1/86,400 of the length of a mean solar day or, simply, a portion of the rotation of the Earth.

For some applications, such as calculating the precise position relative to a satellite orbiting the Earth, we need to convert from the seconds to a rotational

position. One wrinkle is that we can now measure the rotation to high enough precision that slight variations in the rotation matter.

This is all very simple, and while time geeks may debate about some of the details such as how to deal with relativistic effects, such issues are well covered within the scientific community.

Time is used to coordinate independent observations or actions. The very high precision of the atomic clock allows us to look at where a set of satellites are in the sky. It's a modern version of the technique that sailors would use to determine their location using a sextant to see how high a given star is in the sky.

The Human Factor

In early societies, the skies served as an implicit calendar for determining when to plant crops and when to perform rituals.

In a sense, December 21st (or thereabouts) was defined to be the winter solstice so the calendar had to be adjusted to make that true. The problem was that the year had an extra ¼ day.

We tend to use physical metaphors for abstractions so we talk about inserting an extra day (February 29th) or an extra month for lunar calendars but what is really happening is that we are simply adjusting the names of the days to align with the seasons.

The physical metaphors can be useful but they can also cause confusion. When countries shifted from using the naming system of the Julian calendar to the current Gregorian calendar, many people felt as if the time was being taken away rather than simply being renamed.

It's no different from reclassifying Pluto as a minor planet. Changing the category doesn't change the planet itself and renaming October 4th 1582 to October 14th doesn't change where the sun is in the sky.

We have a similar issue with the time of day. Noon, sunset, and sunrise are useful for coordinating local activities with observations of the position of the sun.

With the advantage of modern timekeeping, we could coordinate activities to the precise minute or second. We knew precisely when high noon was and had office hours that could start precisely at 9 AM, mean solar time. Just as the solstice was a defining event for the calendar, the hands on the clock dial represent the position of the sun which has been defining time since the days of the sundial. We've just replaced the shadow with higher-precision pointers.

With the advent of high-speed communications (the telegraph) and high-speed transport (the train), it became important to coordinate with events that might be far away. Instead of using the position of the sun, we accepted an arbitrary scale so that the time called "noon" was the same time in two cities within a time zone. They could share a common reference clock through means of a telegraph line, which allowed the synchronization of trains traveling between the cities.

Daylight savings time shows the power of renaming. Instead of encouraging people to start their workday earlier in the summer, we simply rename 8 AM and call it 9 AM. The fact that we traditionally implement this by moving the hands of a clock gives us another physical metaphor and obscures the fact that we are simply renaming the time and further decoupling it from the position of the sun.

Time zones represent local names. Since GMT was the first time zone, created to coordinate trains in Britain (hence Greenwich Mean Time), it became the base time zone (offset zero). Once one converts local time to GMT, events in different time zones could be coordinated. By using GMT, computers don't have to deal with the crazy quilt of time zones and daylight savings time adjustments.

I'm using the name GMT for the classic time zone without the imposition of the leap second in UTC.

Notation and Computers

Early computers used the same familiar notation with two **columns** each for year, month, day, hour, and minute. We had the Y2K bug because there were only 80 columns and no sense wasting two columns on the century change that wouldn't happen for decades.

Leap seconds are similar. Early computers stored dates the same way with fields for hours and minutes. If astronomers depended on storing dates that way, then the leap second would be necessary to make sure that the minutes were precisely aligned to the rotation of the Earth. There wasn't a place to keep a separate adjustment.

By 1970, computers started to store time internally as simply seconds from the beginning of an epoch, thereby matching the TAI approach of simply counting seconds. The minutes are only used in presenting the date and time to people. This is why Y2K didn't affect navigation computers. And, in the same way, computers that are used to aim telescopes don't depend on the hours and minutes notation. In fact, they can't depend on leap seconds either but rather have to have a finer-grained offset (UT1).

The fact that the second is no longer a fraction of the rotational position of the Earth means that the actual length of the day is a few milliseconds off of the precise 86,400 s. The discrepancy isn't important in the short term for activities that use hours and minutes notation rather than simply counting seconds. In about 5000 years that may become an issue, at which point we can simply do what we do with daylight time and rename the hours.

Lost in Translation

The goal of a leap second is to keep the hours and minutes on wall clock time precisely aligned to the rotational position of the Earth.

Programs that need that precision don't look at the hours and minutes on the wall clock. They use Unix or TAI time counted in seconds. Technically, for tracking satellites, they use UT1 which has continuous fine-grained adjustments based on astronomical observations.

Hours and minutes make it easier for people to coordinate. That coordination isn't tied to the rotation of the Earth. It is more important to make sure the clocks show the same time and adding an extraneous factor like a leap second only makes this more difficult.

The extra precision is not necessary. After all, the wall clock may be off from the sun's position by an entire time zone. Or, in the case of China, by six time zones!

The leap second adjustment is far from benign. It is impossible to convert from seconds to minutes because the definition of a minute is no longer 60 s! For past dates, we might try to keep a table of leap seconds to adjust but that information simply isn't available for future minutes.

Database languages such as SQL can't even represent an irregular minute.

Returning to Simplicity

The simple response to leap seconds should be *don't do that*. After all, it's just a notation and breaking that notation defeats any benefit of the leap second.

We will need an adjustment period as applications that actually do apply the leap second to HMS notation will have to be unwound. We might do that by simply removing a second every 6 months. Applications that don't care about the leap seconds won't be affected and applications that do care will adjust until both groups can share a common conversion of HMS to and from seconds.

Understanding notation and representation are basic to information science. Time isn't simply about the number of minutes or seconds. We might want to specify that an appointment is 4 PM local time independent of the relationship to GMT or that a TV commercial is on the half hour during prime time.

Those descriptions can't be reduced to a simple GMT or TAI time. Rather than focusing on the rotation of the Earth, we must have an articulate understanding of representation if we are to embrace the full power of the new information-based society.

As we've seen the leap second isn't relevant to precise timing which does not rely on using minutes and hours notation. The leap second actually fails at the stated purpose of aligning wall clocks with celestial events because we cannot reliably convert between the two notations for future dates. In fact, we depend on ignoring leap seconds in order to assure that our values in our databases are consistent.

It's time to relegate the leap second to the footnotes of history.

Chapter 34
Common Calendar: Fixed-Epoch Deterministic UTC-Based Local Timescales

Brooks Harris

Abstract A new and comprehensive approach is needed to achieve uniform civil timekeeping across the world. We propose a reference timekeeping framework called common calendar which preserves the age-old tradition of timekeeping by the sun and provides a uniform matrix of fixed-epoch deterministic local timescales. The proposal is made up of seven specifications:

1. *Time-related terms and definitions* provides a comprehensive glossary for the set of specifications, collecting terms from many sources to clarify the use of the UTC (Coordinated Universal Time) standards in general and application to common calendar in particular.
2. *TAI-UTC API* (application programming interface) provides mechanisms for automatic acquisition of TAI-UTC (leap second) history, announcement, and expiration metadata to fill the obvious missing link between UTC time dissemination and the TAI (International Atomic Time) timescale.
3. *Tz database API*—time zone and Daylight Saving Time dynamic metadata.
4. *YMDhms API* details the calculations necessary to perform conversion between seconds and accurate UTC compliant YMDhms representation.
5. *Common calendar local timescales* is at the heart of the proposal, specifying an array of identical fixed-epoch reference local timescales each defined by UTC offsets (time zones) with unambiguous rules for the application of UTC leap seconds, Daylight Saving Time (DST), and YMDhms encoding using the YMDhms API.
6. *Common calendar binary format* defines a binary data format for compact carriage of local timescale date, time, and metadata.
7. *Common calendar character format* provides a comprehensive YMDhms character representation. It augments ISO 8601 recommendations with leap second and DST metadata to provide symmetrical reflection of CBF binary data.

Notation"YMDhms" is shorthand for year-month-day hour:minute:second representation.ISO 8601 representation is supplemented with suffixes (UTC) and (TAI), for example, 1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC)."UTC1972" is shorthand for 1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC).

B. Harris (✉)
EdlMax LLC, New York, NY, USA
e-mail: brooks@edlmax.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_34

A prototype reference implementation in c/c++ has been developed to verify and demonstrate the use of the common calendar specifications. The seven specification documents are in development.

It is hoped common calendar may provide a starting point for formal standardization that finds its way to international acceptance. Comments, ideas, and help improving these specifications are welcomed.

Keywords Timekeeping • UTC • Leap second • Common calendar

Introduction

Timekeeping technologies are good enough for everyday use, but flaws remain that prevent accurate representation of date and time in all cases. Inconsistent treatment of the discontinuities introduced by leap seconds and Daylight Saving Time leads to inconveniences, inaccuracies, errors, and threat of system failure. A new and comprehensive approach is needed to achieve reliable uniform timing across the world.

Many topics of standardization contribute to incompatible system implementation:

- Varied terminology through the literature and the complexity of the UTC standards risks misinterpretation.
- The absence of standardized formatting and dissemination protocols of TAI-UTC information is an obvious missing link that thwarts automated accounting of leap seconds.
- Time zones and Daylight Saving Time rules are not well defined and subject to political policy that may introduce unexpected changes.
- The history of timing specifications has led to systems with mismatched origins and counting methods.
- No standardized algorithm for conversion between seconds and YMDhms representation with native support for leap seconds has been adopted.
- No standardized compact binary data type conveying the necessary and sufficient information for consistent local date and time representation has been proposed.
- Specifications for YMDhms character representation are varied and incomplete.

Common calendar presents a comprehensive design to overcome the ambiguities of interpretation of standards and the indeterminate counting methods of discontinuities.

There are two sets of dynamic metadata that must be available to provide accurate civil timekeeping:

- TAI-UTC (leap seconds) values
- Time zone and Daylight Saving Time rules

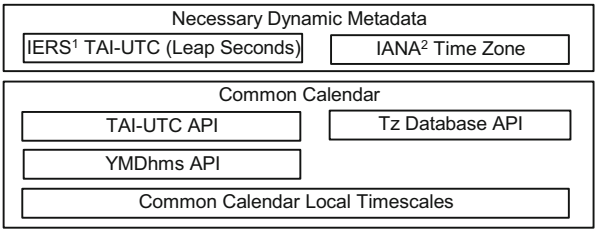
The introduction of leap seconds is unpredictable because of the many factors that affect Earth's rotation. Time zones and the use of Daylight Saving Time are subject to change.

Common calendar addresses these topics by defining APIs to provide access to these two sets of dynamic factors. TAI-UTC API and tz database API aggregate the

dynamic metadata of leap seconds, zones, and DST rules necessary to fully describe local date and time.

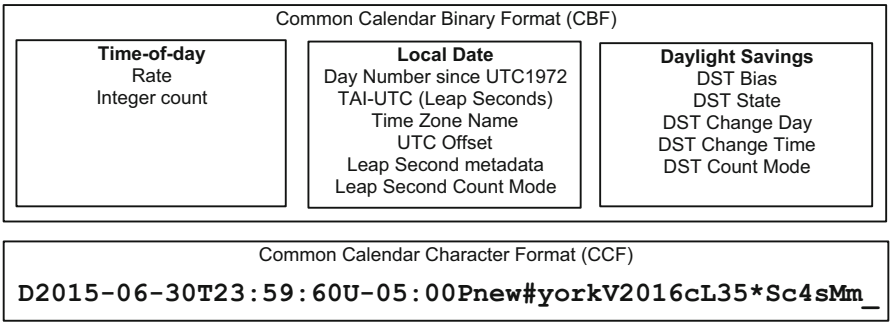
System clocks typically operate on a seconds basis in some form, and conversion between seconds and YMDhms values is fundamental to most implementations. The YMDhms API defines algorithms for these calculations with native support for leap seconds as interfaced to the TAI-UTC API.

Together these APIs collect all the necessary raw information into normalized data representations and algorithms to represent local time in YMDhms form. The common calendar local timescales define a matrix of local timescales consistent with these APIs, with counting methods, rules, and policies to enable deterministic civil timekeeping.



IERS[1] IANA[2]

Common calendar provides two formats of data representation for internal interfacing, interchange, and input and output. Common calendar binary format (CBF) provides a compact binary data form suitable for fast transfer, efficient interchange, and economical storage. Common calendar character format (CCF) is machine-readable ASCII format in YMDhms form with comprehensive metadata. The two formats are symmetrically convertible to one another using the YMDhms API methods.



[1]IERS: International Earth Rotation and Reference Systems Service.

[2]IANA: Internet Assigned Numbers Authority.

Common calendar presents seven tightly coupled specifications which together overcome the ambiguities of standard interpretation and indeterminate accounting of YMDhms discontinuities:

- **Date- and Time-Related Terms and Definitions**—comprehensive collection of terms and definitions to clarify the use of UTC and local time and provide common lexicon for the common calendar specifications

  For dynamic metadata—leap seconds, time zones, and Daylight Saving Time:

  - **TAI-UTC API**—data types and operations for automated access to dynamic IERS TAI-UTC information
  - **Tz Database API**—data types and operations for automated access to dynamic IANA Time Zone Database

- **YMDhms API**—data types and algorithms for conversion between seconds and YMDhms representation with native support of leap seconds
- **Common Calendar Local Timescales**—defines a matrix of symmetric local timescales each defined as a UTC offset from the UTC1972 origin and UTC-00:00
- **Common Calendar Binary Format**—compact binary data format to facilitate fast transfer, efficient interchange, and economical storage
- **Common Calendar Character Format**—ASCII character-based machine-readable representation of date and time in a YMDhms form to impart familiar meaning to human users

These specifications have been prototyped and tested in a C++ implementation. The seven specification documents are in development.

Each component is described in the following sections.

Date- and Time-Related Terms and Definitions

Any standard requires a comprehensive terms and definition specification. This is difficult in timekeeping because so many terms have been used, and misused, through history. The specifications of UTC are scattered through many standard organizations using various nomenclatures leading to some confusion and implementation inconsistency.

To address this challenge, common calendar has collected terminology in the *date- and time-related terms and definitions* document. It consolidates the descriptions of TAI and UTC, collecting the many details of the dispersed international standards and clarifying the chain of provenance among those specifications. The basic timekeeping definitions are drawn from ISO 8601, its underlying IEC specifications, the *BIPM brochure* (Bureau International des Poids et Mesures, the International System of Units (SI)) and the *BIPM International Vocabulary of*

Metrology (VIM). This includes description of the origins of TAI and UTC and the counting methods of the application of TAI-UTC (leap seconds) to the UTC timescale.

One aim of this document is to draw out and clarify some basic facts of UTC:

- 1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC) is the calibration point at which the relationship between atomic time and observed mean solar time was made to converge on exactly 10 integral seconds as determined by the development process of UTC.
- 1958-01-01 00:00:00 (TAI) is the origin of the TAI timescale. This origin had been used previously, but the choice was retroactively affirmed and given a more precise definition and value with respect to the new UTC calibration point, 441,763,200 s before 1972-01-01 00:00:00 (TAI) or 441,763,210 s before 1972-01-01 00:00:00 (UTC).
- Leap seconds are integral one-second adjustments decided and announced by the IERS to keep the difference between TAI and UTC less than 0.9 s.
- A leap second occurs in the UTC count at the last second of the day *before* the date on which the IERS has declared the leap second. A positive leap second inserts an additional second into the count (23:59:60 added); thus that day is 86,401 s duration. A negative leap second would omit, or "drop," a second count (23:59:59 omitted) at the end of the day. A negative leap second has never happened, but applications should be prepared to account for the resulting 86,399 second day and rollover to the next day at 23:59:58.
- The TAI-UTC history defines the difference between the TAI and UTC timescales, and this information is required to calculate durations between two time points on the UTC timescale.

Dynamic Metadata: Leap Seconds, Time Zones, and Daylight Saving Time

There are two sets of dynamic metadata that must be available to provide the necessary and sufficient information for accurate civil timekeeping:

- TAI-UTC (leap seconds) values
- Time zone and Daylight Saving Time rules

The introduction of leap seconds is unpredictable because of the many factors that affect Earth's rotation. Time zones and observation of Daylight Saving Time are politically and culturally selected and subject to change. This necessitates timely access to updated information for both factors.

This unavoidable unpredictability limits the range of possible prediction. No timekeeping system can accurately calculate a future date time beyond either the expiration date of the current leap second information or the current valid time zone and DST rules, whichever is earlier. Calculations in the past require access to the

full history of leap seconds, time zone changes, and DST occurrences to represent local time across the full extent of the timescale.

Common calendar addresses these topics by defining APIs to provide access to these two dynamic factors. TAI-UTC API and tz database API aggregate the dynamic metadata of leap seconds, zones, and DST rules necessary to fully describe local date and time.

TAI-UTC API: Leap Second Dynamic Metadata

TAI-UTC, the integral second difference between TAI and UTC, is at the foundation of modern civil timekeeping. Obtaining and maintaining this information is critical to accurate timekeeping, but no formal mechanism for automatic access to this metadata has been adopted. The TAI-UTC API addresses this obvious missing link in timekeeping technologies.

The IERS issues Bulletin C "either to announce a time step in UTC or to confirm that there will be no time step at the next possible date." Bulletin C was originally intended as a human-readable publication requiring operator intervention in a timekeeping system to make appropriate UTC adjustments.

Today the IERS publishes Bulletin C in several forms including text and XML versions (http://www.iers.org/IERS/EN/DataProducts/EarthOrientationData/eop.html). They also provide the leap second web service (http://hpiers.obspm.fr/eoppc/index.php?index=webservice). One IERS publication, Leap_Second_History.dat (https://hpiers.obspm.fr/eoppc/bul/bulc/Leap_Second_History.dat), provides a concise tabulation of the previous leap second introductions including the state of the current Bulletin C.

All these IERS resources are helpful, but none provide a comprehensive and uniformly formatted version of the leap second information suitable for use in an automated dissemination mechanism.

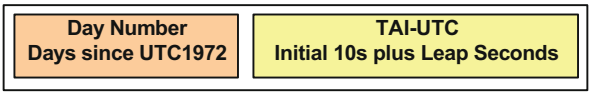
The TAI-UTC API provides methods for automatic dissemination and acquisition of TAI-UTC metadata to fill the obvious missing link between UTC time dissemination and the TAI timescale. TAI-UTC API defines data types and required operations for uniform treatment of the TAI-UTC (leap seconds) history, announcement, and the expiration date of valid leap second information. The API's goal is to provide uniform timekeeping results through any technology employed for leap second metadata dissemination and consistent application of the data to timekeeping systems. It is applicable to existing timekeeping technologies and timescales and crucial to any implementation of common calendar.

The proposed TAI-UTC data format is based on a packed binary pair consisting of a 21-bit day number and an 11-bit TAI-UTC offset. This provides a day and TAI-UTC counter in a 32-bit data structure with a range of approximately 3000

years.[3] With this compact binary data representation, each TAI-UTC value is small, and it is then possible to construct an array of the TAI-UTC history (a leap second table) of reasonable size.

The TAI-UTC API defines day 0 to be UTC1972 (1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC)). This corresponds with MJD 41317 as used by the IERS in https://hpiers.obspm.fr/iers/bul/bulc/Leap_Second_History.dat.

This leads to the definition of the TAI-UTC API date data element:



TAI-UTC API date data element

An array of TAI-UTC API data elements is used to construct a table of TAI-UTC history corresponding to the IERS Leap_Second_History.dat information. An additional TAI-UTC API data element, called EXPIRATION, is appended as the last element of the array to signal the expiration date of the validity of the last TAI-UTC value.

The next-to-last element always holds the most recently announced leap second, and the TAI-UTC value of EXPIRATION is the same as the TAI-UTC value in that next-to-last element. When the IERS announces a leap second, the announced values are inserted as a new element in the next-to-last position, and the values of the EXPIRATION element are updated.

The expiration date can be obtained from the IERS Leap_Second_History.dat file. It includes a line in the comment section stating the file's "expiration" date, for example, `# File expires on 30 June 2016.`

It is not an official obligation of IERS to provide this expiration date. For purposes of the TAI-UTC API, the expiration date is taken to be a declaration by IERS that no leap second will be declared before that date.

IERS Leap_Second_History.dat supplies the date of leap second introductions in the form of modified Julian date (MJD). The IERS Leap_Second_History.dat data is interpreted by TAI-UTC API as:

[3]TAI-UTC data format is said to support a range of "3000 years." A rough approximation of the range of a 21-bit day counter (MAX 2097152) is 2097152/(365.25) = 5741.689254 years. The number of leap seconds that may occur is unknown and there are many methods of estimates in the literature. A straight extrapolation of the average as known from the past is 2016 – 1972 = 44 years/(36 – 10 = 26 leap seconds) = ~1.6923 years per leap second. The TAI-UTC range of 11 bits (MAX 2048) gives an estimate of 1.6923 × 2048 = ~3465.8461 years. The TAI-UTC counter range is the limiting factor of the total range. Some projections suggest the rate of leap seconds will likely increase over centuries. Rounding 3465.8461 years down to "approximately 3000 years" is used as a rough claim of the supported TAI-UTC API range (1972 + 3000 = year 4972).

TAI-UTC API date data element

Element name	Description	Type	Range
Day number	Zero-based count of days since 1972-01-01T00:00:00 (UTC) = MJD as published by IERS minus 41317	Integer	0 to 2097152-1
TAI-UTC value	TAI-UTC value as published by IERS	Integer	0 to 2048-1

- MJD 41317 = UTC1972 (1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC))
- MJD 41317 time-of-day seconds = 0 (zero) = midnight
- MJD length of day = 86,400 s
- TAI-UTC values include the UTC1972 10 s initial calibration plus accumulated leap seconds

Tz Database API: Time Zone and Daylight Saving Time Dynamic Metadata

"Civil time," often called "local time," refers to timescales established by law or custom in some jurisdiction at some geographic location for legal, commercial, and social purposes.[4]

Most civil time is based on UTC, usually defined by its offset from UTC and any Daylight Saving Time (DST) rules in use. The details of how these rules describe local time are typically guided by several standards, de facto standards, and common practice.

Unfortunately there is no universally applicable set of formulas that strictly define local time. The rules cannot be 100% correct in all places in all circumstances because some jurisdictions may define local time in some unusual way or make sudden changes. Nonetheless, common practice has led to a high degree of uniformity, accuracy, and confidence in most places where responsible decisions are made. It is especially important that notice of any upcoming changes is given

[4]There are generally two types of time zone in use: civil (land) and nautical. Civil time zones are usually designated as a time offset from the UTC applicable to some territory on land. Nautical time zones are specified by longitude for purposes of navigation at sea. Common calendar is concerned only with civil time and does not address nautical time zones (see *Date and Time Terms and Definitions, 11 Civil Time (Local Time)*).

well in advance, and this has typically been the case in responsibly managed jurisdictions.[5]

IANA maintains the Time Zone Database (https://www.iana.org/time-zones). It is the most comprehensive source of time zone and DST information. Originally known as the "Olson database" or "tz database," it has been the de facto standard for time zones and DST rules for many timekeeping systems for decades. It has recently gained further stature since IANA has taken on responsibility for its oversight and management.

The zones and rules as known to the IANA Time Zone Database system are made available in the "time zone data" files tzdata.tar.gz and are available as described in the "Distribution" section of the Web site. The zones and rules for Daylight Saving Time are maintained as described by IETF RFC (Internet Engineering Task Force Request for Comments) 6557 (http://tools.ietf.org/html/rfc6557).

IETF RFC 7808, time zone data distribution service (https://tools.ietf.org/html/rfc7808), has recently been completed, lending new interoperability flexibility to managing time zones and DST rules.

Common calendar tz database API is an interface to the IANA Time Zone Database data to manage this critical set of dynamic metadata. It adapts data types for zones and rules in a form and manner consistent with common calendar local timescales.

Tz database API conforms to the IANA Time Zone Database rules governing the zone and DST data and interprets their values using the formulas of common calendar local timescales and YMDhms API. This approach bypasses the POSIX-based tzcode c language implementation in favor of reading the local time rule parameters without reliance on any date and time calculations. Subsequent processing applies the YMDhms API formulae with its native support of leap seconds. The common calendar c++ reference implementation details these operations.

Common calendar and tz database API enhance the IANA Time Zone Database system by adopting zone and DST data update policies to allow deterministic prediction for at least 4 months. Common calendar binary format and character format record the IANA Time Zone Database tzdata release version in their metadata to provide an audit trail for forensic determination of local time if required.

[5]For example, the US Energy Policy Act of 2005 was signed into law on August 8, 2005, specifying a new rule for DST onset as the first Sunday on or after the 8th of March. This took effect Sunday March 11, 2007, providing 22 months of advance notice.

YMDhms API: Seconds to YMDhms Conversion

The main objective of any civil timekeeping system is to produce date and time in the familiar year, month, day, hour, minute, second (YMDhms) form suitable for human understanding. Systems typically operate internally on a seconds basis, and conversion algorithms between seconds and YMDhms values are fundamental to most implementations.

Many systems adopt the classic operations, or functions, described by POSIX and named mktime() (YMDhms-to-seconds) and gmtime() (seconds-to-YMDhms). These execute the Gregorian calendar counting method, including its leap years, making calculations based on 86,400 second day durations. POSIX has recognized limitations in supporting UTC and leap seconds; in particular it cannot count to 23:59:60 and has no mechanism to obtain and apply the TAI-UTC leap second values.

To overcome this deficiency, common calendar specifies the YMDhms API. It specifies data types and operations to perform the conversion between seconds and YMDhms with leap second compensation using the TAI-UTC API for access to the required dynamic TAI-UTC leap second information.

The YMDhms API specifies data types analogous to POSIX time_t ("seconds since the Epoch") and struct tm ("broken-down time"), called i641972Time_t (signed 64-bit seconds since UTC1972) and YMDhms_tm (YMDhms, leap seconds, and additional metadata). YMDhms API describes functions analogous to POSIX mktime() and gmtime() calling them YMDhmsToSeconds() and SecondsToYMDhms().

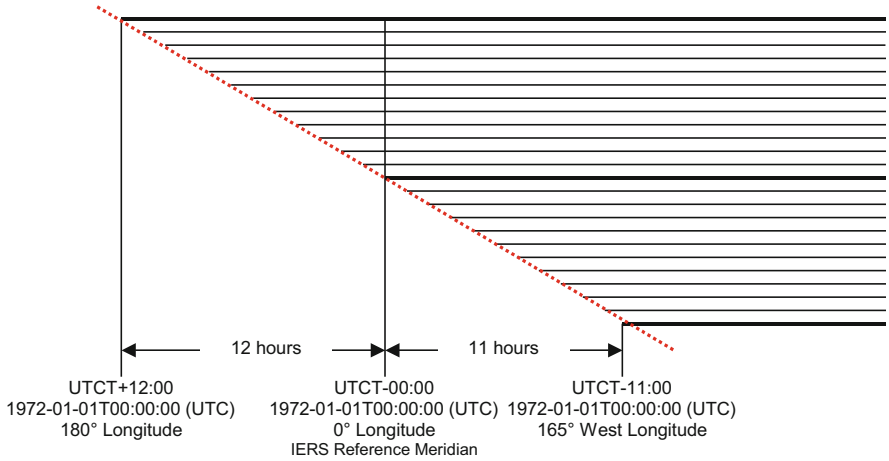
The POSIX formulas act on a UTC1970 epoch, while YMDhms API acts on the same origin as the TAI-UTC API: UTC1972 (1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC)). YMDhms_tm includes additional information to POSIX struct tm, including TAI-UTC values, and directly supports leap second compensation including count to 23:59:60 for positive leap seconds and 23:59:58 rollover for negative leap seconds.

The c++ reference implementation includes the explicit algorithmic operations.

Common Calendar Local Timescales

Common calendar is made up of a set of 105 local timescales each defined as a fixed time offset from UTC time. These are arranged in 15-min segments spanning the 26 h needed to account for existing traditional "time zones" as inventoried by the IANA Time Zone Database.[6] Unique names for each local timescale reflect their

[6]Local timescales 0–104 (105) covering existing time zones, 12 negative signed, 1 for −00:00 at UTC, 12 positive signed plus 2 positive signed for +13:00 (Pacific/Tongatapu) and +14:00 (Pacific/Kiritimati). See Annex - Common Calendar Local Timescales.



Common calendar local timescales definitions

UTC time offsets (e.g., "UTCT-05:00"[7]) to avoid the ambiguity encountered in traditional naming conventions. Local timescales are similar to conventional "time zone" timescales but defined as separate standardized local timescales in the time domain. They can be mapped to conventional time zones using the tz database API. UUIDs are assigned to ensure unique identity in digital environments.

The YMDhms counting method of the local timescales is the Gregorian calendar augmented with carefully defined compensation for UTC leap seconds. The origin of UTC time is UTC1972 and this is the reference origin for all components of common calendar. Each local timescale is an independent timescale defined by its fixed time offset from UTC time, and each has an origin labeled 1972-01-01 00:00:00 (UTC).

Common calendar specifies the use of an idealized Euclidean geometry sphere to approximately represent the Earth and thus approximate longitudes for purposes of civil timekeeping. 0° longitude corresponds to the IERS reference meridian (IRM). Each local timescale UTC offset is associated with a longitude position.

Conventional time zones and Daylight Saving Time are mapped to the local timescales using the tz database API. Local timescales are defined by their fixed time offset from UTC time, and this does not change except in the rare case where a time zone's UTC offset is actually changed. Daylight Saving Time is applied to the local timescales as an adjustment to the date-time value on that timescale, not as shifts to another local timescale. This approach eliminates the uncertainty introduced by the common practice of shifting the UTC offset to apply DST.

[7]"UTCT" is used in the character string name of the UTC offset to differentiate it from other uses of "UTC."

Two DST "count modes" are defined for YMDhms representations: "conventional" and "uninterrupted." "DST conventional count mode" produces the familiar and expected discontinuity in the YMDhms sequence with DST onset or retreat. "DST uninterrupted count mode" is provided for systems or applications that cannot tolerate such a discontinuity. This count mode delays the DST shift in the YMDhms sequence until the end of the day yielding an uninterrupted incrementing YMDhms value until midnight. This produces a 23-hour day on DST onset days and a 25-hour day for DST retreat days for the typical 1 h DST bias. This requires the seconds-to-YMDhms algorithm to have support counting to 24:59:59. This capability is defined in the YMDhms API and shown in the c++ reference implementation. See Annex - c++ Listings, Listing - Common Calendar Binary Format, excerpts from CBF.h, CBFDstCountMode_et.

Common calendar specifies a policy of recording the local date and time as known to the emitting system in all cases. This brings clarity and specificity to the meaning of the data and consistency and simplicity to generating and interpreting common calendar date and time values.

Leap seconds are treated as permanent members of the timescales, while Daylight Saving Time is treated as transient modifiers of the YMDhms encoding.

Time-of-Day (TOD) Count Mode and Leap Second Introduction

Common calendar employs a counting method called time-of-day (TOD) count mode to define the relation of date and time-of-day portions of the YMDhms representation of the underlying seconds-since-1972 count:

- **Leap Second at UTC**—leap seconds are introduced in the local timescale YMDhms sequence simultaneous with introduction on the UTC timescale. This supports the common practice employed in many systems and time dissemination protocols.
- **Leap Second at Midnight**—leap seconds are introduced in the local timescale YMDhms sequence at midnight. This variation on common practice produces a simplified and symmetrical matrix of local time representations to support systems that must avoid discontinuities in the YMDhms counting sequence.
- **None**—there is no relation between the date and time-of-day: the hours-minutes-seconds portion of the YMDhms representation is independent of the year-month-day portion. This supports representation of simple running clock values obtained on a given date, like a "stop watch" or "game clock."

The leap second at UTC and leap second at midnight TOD count modes are further described in the following sections.

TOD Count Mode: Leap Second at UTC

Midnight on any local timescale is defined by its YMDhms representation of UTC offset and Daylight Saving shifts, if used. Many timekeeping systems adopt the common practice of introducing leap seconds on local timescales simultaneous with their introduction on the UTC timescale.

Unfortunately this generates a discontinuity in the YMDhms counting sequence sometime during the day between the local midnights (midday), and this produces a complex and asymmetrical relationship among the local timescales during leap second days. Despite its being difficult to understand, implement, and test, this convention is supported by common calendar to provide compatibility with these widely deployed systems.

Common calendar supports this conventional leap second counting scheme in three variations:

1. TOD_LEAPSECOND_UTC_UTC—leap seconds are introduced simultaneous with UTC on local timescales, and the leap second is labeled **xx:59:60** as per UTC specification.

   For example[8]:

Time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/ Berlin	`D2015-07-01T01:59:59U +01:00PberlinV2016cL35+Sc4sMm_`	1372640424
Europe/ Berlin	`D2015-07-01T`**`01:59:60`**`U +01:00PberlinV2016cL35*Sc4sMm_`	1372640425
Europe/ Berlin	`D2015-07-01T02:00:00U +01:00PberlinV2016cL36+Sc4sMm_`	1372640426

2. TOD_LEAPSECOND_UTC_NTP—leap seconds are introduced simultaneous with UTC on local timescales, and the leap second is labeled **xx:59:59** as per NTP specification (“freeze”).

   For example:

Time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/ Berlin	`D2015-07-01T01:59:59U +01:00PberlinV2016cL35+Sc4sMm_`	1372640424
Europe/ Berlin	`D2015-07-01T`**`01:59:59`**`U +01:00PberlinV2016cL35*Sc4sMn_`	1372640425
Europe/ Berlin	`D2015-07-01T02:00:00U +01:00PberlinV2016cL36+Sc4sMm_`	1372640426

[8]Please see the section “Common Calendar Character Format” (CCF) for details of character representation.

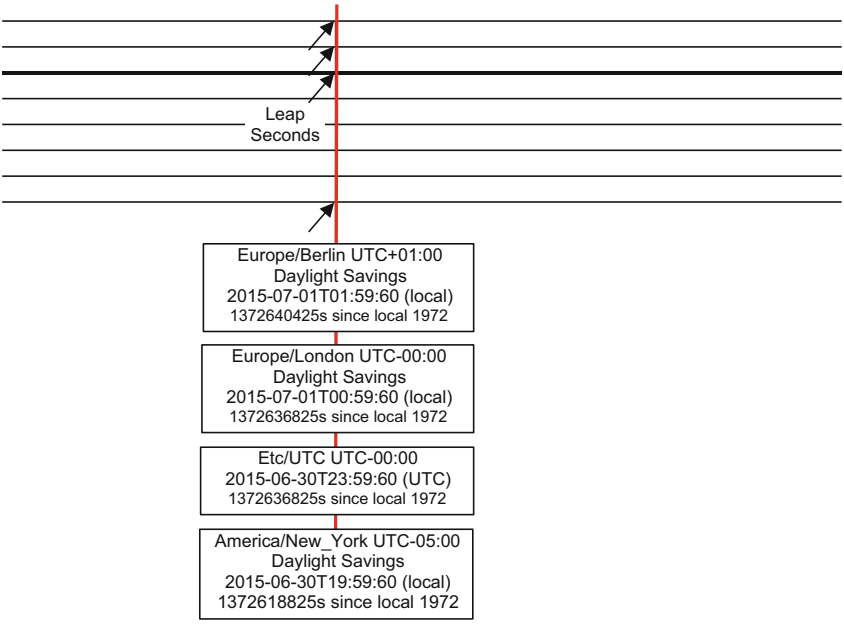
3. TOD_LEAPSECOND_UTC_POSIX—leap seconds are introduced simultaneous with UTC on local timescales, and the leap second is labeled **xx:00:00** as per POSIX specification ("rollover and reset"). For example:

Time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/ Berlin	D2015-07-01T01:59:59U +01:00PberlinV2016cL35+Sc4sMm_	1372640424
Europe/ Berlin	D2015-07-01T**02:00:00**U +01:00PberlinV2016cL35*Sc4sMp_	1372640425
Europe/ Berlin	D2015-07-01T02:00:00U +01:00PberlinV2016cL36+Sc4sMm_	1372640426

The following figure and table shows examples of selected time zones for the 2015 leap second in TOD_LEAPSECOND_UTC_UTC count mode.

TOD Count Mode: Leap Second at Midnight

The common calendar "leap second at midnight" counting scheme introduces the leap second at the end of a leap second day in the local YMDhms representation to avoid the midday discontinuity produced by the conventional "leap second at UTC" method. This results in a logically consistent and symmetrical matrix of local timescales, each with identical leap second counting methods and uniform phase,



TOD_LEAPSECOND_UTC_UTC count mode—2015 leap second introduction

TOD_LEAPSECOND_UTC_UTC count mode—2015 leap second introduction

Leap second time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/Berlin	D2015-07-01T01:59:60U +01:00PberlinV2016cL35*Sc4sMu_	1372640425
Europe/London	D2015-07-01T00:59:60U-00:00PlondonV2016cL35*Sc4sMu_	1372636825
Etc/UTC	D2015-06-30T23:59:60U-00:00PetcutcV2016cL35*Mu_	1372636825
America/New_York	D2015-06-30T19:59:60U-05:00Pnew#yorkV2016cL35*Sc4sMu_	1372618825

duration, and offset relationships among them. This simplifies implementation and testing which can lead to more consistent local time stamp representation.

The "leap second at midnight" counting scheme was designed to support timekeeping systems that cannot tolerate midday YMDhms discontinuities. It may be chosen by systems seeking to simplify timekeeping procedures for more consistent and deterministic local time stamp implementation.

TOD_LEAPSECOND_MIDNIGHT count mode introduces leap seconds labeled **23:59:60** at the end of a leap second day on each local timescale. For example:

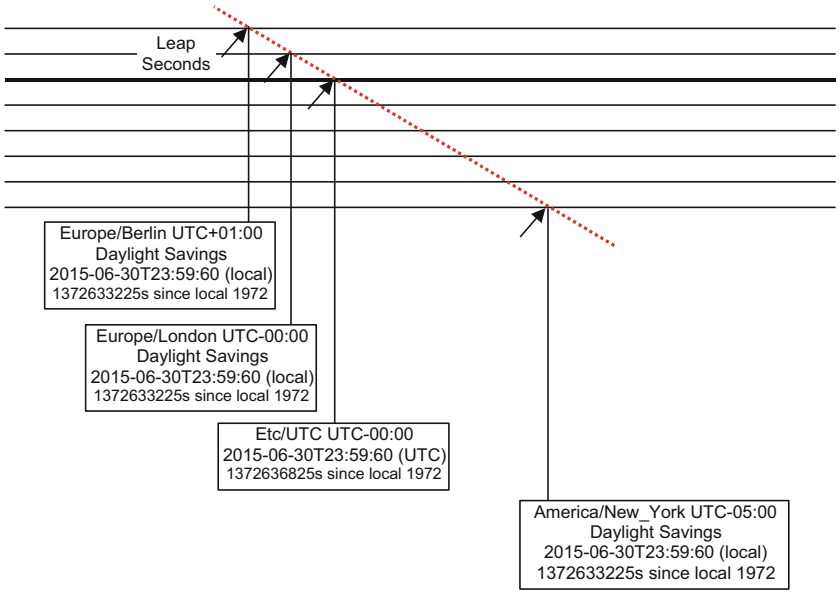
Time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/Berlin	D2015-06-30T23:59:59U +01:00PberlinV2016cL35+Sc4sMm_	1372633224
Europe/Berlin	D2015-06-30T**23:59:60**U +01:00PberlinV2016cL35*Sc4sMm_	1372633225
Europe/Berlin	D2015-07-01T00:00:00U +01:00PberlinV2016cL36Sc4sMm_	1372633226

This count mode eliminates the midday discontinuities introduced in YMDhms representations by the common practice of introducing leap seconds simultaneous with its appearance in UTC. Instead, it places the leap second introduction at the end of a leap second day on local timescales consistent with the UTU-R Rec 460 specification.

The following figure and table shows examples of selected time zones for the 2015 leap second in TOD_LEAPSECOND_MIDNIGHT count mode.

Compatibility to NTP, PTP, POSIX, and GPS

There are several established time dissemination systems widely used in civil timekeeping: NTP, PTP, POSIX, and GPS. To facilitate interoperability with



TOD_LEAPSECOND_MIDNIGHT count mode—2015 leap second introduction

TOD_LEAPSECOND_MIDNIGHT count mode—2015 leap second introduction

Leap second time zone	YMDhms character representationcommon calendar character format (CCF)	Seconds since local 1972
Europe/Berlin	`D2015-06-30T23:59:60U +01:00PberlinV2016cL35*Sc4sMm_`	1372633225
Europe/London	`D2015-06-30T23:59:60U-00:00PlondonV2016cL35*Sc4sMm_`	1372633225
Etc/UTC	`D2015-06-30T23:59:60U-00:00PetcutcV2016cL35*Mm_`	1372636825
America/New_York	`D2015-06-30T23:59:60U-05:00Pnew#yorkV2016cL35*Sc4sMm_`	1372633225

these systems, common calendar timescale defines a set of unique names of offsets from the common calendar native UTC1972 origin. Each defines an exact integral second offset to the origins of these timescales.

The local timescales can be converted to existing and traditional timescales. However, the reverse conversion of traditional to CCT local timescales may not yield deterministic results in all cases, in particular at time points of leap second introduction or when local time metadata is unavailable, out of date, or insufficient.

Origins of established timescales

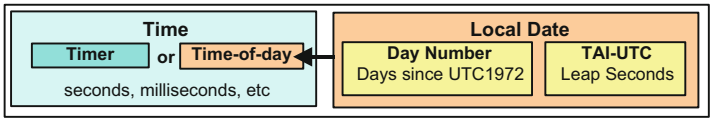
Origin name	Seconds since UTC1972	Coincides with origin of	UTC	TAI-UTC	TAI
UTC1900	−2272060800	NTP Era 0	1900-01-01T00:00:00 (UTC)	10	1900-01-01 00:00:10 (TAI)
TAI1958	−441763210	TAI	1957-12-31T23:59:50 (UTC)	10	1958-01-01 00:00:00 (TAI)
TAI1970	−63072010	1588/PTP	1969-12-31T23:59:50 (UTC)	10	1970-01-01 00:00:00 (TAI)
UTC1970	−63072000	POSIX	1970-01-01T00:00:00 (UTC)	10	1970-01-01 00:00:10 (TAI)
TAI1972	−10	UTC + 10s	1971-12-31T23:59:50 (UTC)	10	1972-01-01 00:00:00 (TAI)
UTC1972	0	UTC	1972-01-01T00:00:00 (UTC)	10	1972-01-01 00:00:10 (TAI)
UTC1972_7_1	+15724800	First Leap	1972-07-01T00:00:00 (UTC)	11	1972-07-01 00:00:11 (TAI)
UTC1980_1_6	+252892800	GPS Epoch	1980-01-06T00:00:00 (UTC)	19	1980-01-06 00:00:19 (TAI)

Common Calendar Binary Format

Common calendar binary format (CBF) provides a compact binary data format to facilitate fast transfer, efficient interchange, and economical storage. The CBF design is intended for binary systems, protocols, and languages, such as embedded systems, time dissemination, and c/c++ while offering compatible expression on other platforms and formats such as Java and XML.

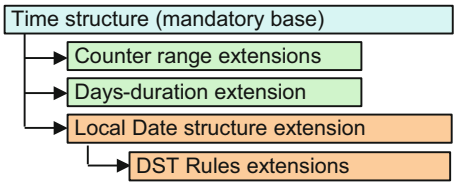
CBF acts in concert with common calendar character format (CCF) to provide comprehensive description of local date and time with symmetrical conversion between the binary and character-based YMDhms representations.

The CBF binary time stamp is a variable length compound counter with associated metadata. It is made up of the mandatory base time structure that provides the anchor for any configuration of CBF, and optional extensions, including the important local date structure.



The variable size design provides a single syntax to construct binary time stamps to accommodate use as a simple 24-hour timer (using only the mandatory time structure) or a full local date and time-of-day time stamp (with the local date

extension). This flexibility minimizes the total size of the time stamp for each purpose.



The mandatory time structure contains a 32-bit unsigned counter and sign bit (signed 33 bit). The range may be increased to 48 or 64 bits using the counter extensions. It can represent resolutions from seconds to picoseconds depending of the rate indicator. Used by itself it is a simple 24-hour timer.

The local date structure extension is added to provide full local date and time-of-day representation. Its calendar date uses the TAI-UTC API date data structure which includes the day-number-since-UTC1972 and TAI-UTC (leap seconds). Its time zone metadata uses elements derived from the tz database API, and its local UTC offset conforms to the common calendar local timescales specification.

Additional leap second metadata provides sufficient information for interpretation by a simple receiving application that has no leap second or tz database information available and facilitates conversion to the CCF YMDhms character representation.

When local date is used, the values of the base time structure are interpreted as time-of-day. Together, the compound counter, made up of the time-of-day from time, and the day-number and TAI-UTC from local date represent the local date and time as an absolute seconds-since-UTC1972 value. For example, at a seconds resolution (rate), the formula is:

```
seconds-since-UTC1972=(DayNumber*86400)+(TAIUTC-10)+(seconds-
since-midnight)
```

The design, with its compound data elements, lends advantage to reading and interpreting a common calendar time stamp. The emitting application has the unavoidable responsibility to accurately populate the time stamp parameters, while it has access to leap second and tz database information. Important aspects of the emitter's processing are transferred to the time stamp data and metadata, and this simplifies the receiver's role. The receiver is relieved from performing some computationally intensive processing, and there is no need for access to leap second and local time information.

Common calendar specifies a policy of recording the local date and time as known to the emitting system in all cases. This brings clarity and specificity to the meaning of the data and consistency and simplicity to generating and interpreting common calendar time stamp values. This allows direct comparison of local time

stamps without the need for the usual methods of converting between "local time" and "GMT time."

An outline CBF features:

1. **Time Structure**—mandatory fixed size 48-bit structure containing:
   - Clock rate enumeration (supports seconds, milliseconds, nanoseconds, etc.)
   - 32-bit counter
   - Counter size extensible to 48 and 64 bits
   - Days-duration extension supports use as 179-year duration counter
   - Flag to signal the presence of local date structure extension
2. **Local Date Structure**—optional fixed size 64-bit local date structure containing:
   - The CCT origin is UTC1972 (1972-01-01 00:00:10 (TAI) = 1972-01-01T00:00:00 (UTC)).
   - CCT always records local time as known to the emitting application.
   - Ordinal day counter with UTC-TAI (leap seconds) for approximately 3000-year range.
   - UTC offset local timescale designation.
   - IANA tz database zone.
   - DST bias, change time, and state.
   - Time-of-day (TOD) count mode defines the linkage between the time and local date:

     TOD_LEAPSECOND_UTC_UTC—leap seconds are introduced simultaneous with UTC on local timescales labeled as 23:59:60 as per UTC specification.

     TOD_LEAPSECOND_UTC_NTP—leap seconds are introduced simultaneous with UTC on local timescales labeled as xx:59:59 ("freeze") as per NTP specification.

     TOD_LEAPSECOND_UTC_POSIX—leap seconds are introduced simultaneous with UTC on local timescales labeled as xx:00:00 ("rollover and reset") as per POSIX specification.

     TOD_LEAPSECOND_MIDNIGHT—leap seconds are introduced "at midnight" on each local timescale labeled as 23:59:60 as per UTC specification.

     TOD_86400S_DAY_DATE—date and time-of-day data are treated as 86,400 second days (leap seconds are unavailable or unknown).

     TOD_NONE—time not related to date.

   - Flag to signal the presence of the optional leap second interval extension with previous and next leap second or expiration.
   - Flag to signal the presence of the optional DST schedule extension with previous and next DST change date and time.

Common Calendar Character Format

Common calendar character format (CCF) is a character-based machine-readable representation of date and time in a YMDhms form to impart familiar meaning to human users. It is complete and comprehensive to provide accurate and deterministic representation and symmetrical conversion to binary CBF.

CCF starts with the guidelines set out in ISO 8601 and adds important variations. The 8601 scheme is augmented in several ways including:

1. Formatting is constrained to YYYY-MM-DD hh:mm:ss form.
2. Clock rate is signaled.
3. Local time is designated as a *fixed* UTC offset (this does not change to indicate DST shifts).
4. Time zone and DST metadata from IANA Time Zone Database is added.
5. TAI-UTC (leap second) value is added.
6. Daylight Saving metadata is added.
7. The 8601 "Z" is not used.

CCF is divided in two main sections: date and time, reflecting the main ISO 8601 organization and the CBF binary structure's division into local date and time. Each may be used independently or together to form a complete date and time-of-day:

1. Time only—represents 24 h of time with no relation to time-of-day, like a "stop watch"
2. Date only—represents the midnight of a calendar day, that is, YYYY-MM-DD 00:00:00 (UTC)
3. Date and time—represents local calendar day and time-of-day

CCF is encoded in several elements delimited by single designated uppercase letters to facilitate parsing. The syntax is variable length with no white space.

Designated characters following the seconds field encode the decimal fractions of seconds-enumerated clock rates.

CCF element delimiters

Delimiting character	Element	Example
`T`	Time or time-of-day	`T01-00:00.999999999_`
`D`	Local time (date in local time zone)	`D2016-02-22`
`U`	UTC offset	`U-05:00`
`P`	Time zone (place)	`Pnew#york`
`V`	Tz database data release version	`V2016c`
`L`	TAI-UTC offset (leap seconds)	`L36`
`S`	Daylight Saving bias and state	`Sc33`
`M`	Time-of-day (TOD) count mode	`Mm`
`_`	Hard terminator	

CCF clock rate enumerations

Enumeration	Rate	Character
CLOCK_0	Seconds	
CLOCK_1	10/1 seconds	t
CLOCK_2	100/1 seconds	h
CLOCK_3	Milliseconds	m
CLOCK_6	Microseconds	u
CLOCK_9	Nanoseconds	n
CLOCK_12	Picoseconds	p
CLOCK_15	Femtoseconds	f
CLOCK_18	Attoseconds	a
CLOCK_21	Zeptoseconds	z
CLOCK_24	Yoctoseconds	y
CLOCK_44	Planck time	k

CCF Examples:

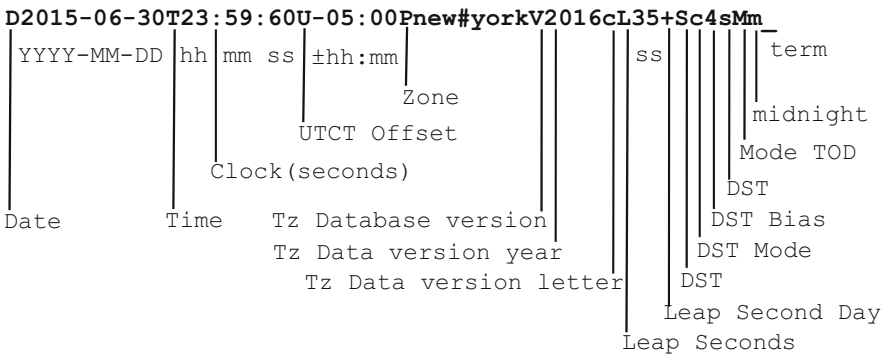
Seconds

```
T01:00:00_
|hh|mm ss term
|   Clock (Seconds)
Time
```

Milliseconds

```
T01:00:00m999_
|hh|mm ss|ddd_term
|   Clock Milliseconds
Time
```

2015 Leap Second, New York



Discussion

Common calendar has been developed to resolve the ambiguities and inconsistencies in current timekeeping specifications and technologies to realize a practical engineering solution for accurate local timekeeping. This has been motivated by the author's work at the Society of Motion Picture and Television Engineers (SMPTE) where reverse compatibility to existing media timekeeping technologies is paramount.

The existing SMPTE time code on which the industry depends was not designed to support truly accurate time-of-day and does not provide means of representing accurate year-month-day information. The infrastructure of storage, transport protocols, and applications that rely on it cannot tolerate discontinuities in the hours-minutes-seconds-frames sequence.

SMPTE has recently developed a new specification for synchronization over IP networks called SMPTE ST2059. ST2059 is based on IEEE 1588/PTP and seeks to align the video signal and time code to the fixed 1588/PTP epoch and generate accurate local time-of-day labeling of video signals. To achieve this requires a rigorous implementation of UTC and local time parameters including Daylight Saving Time. It is here the video industry encounters the challenges of consistent implementation and representation of UTC and local time.

Common calendar is the result of the author's work in meeting that challenge. An initial design goal was to provide a YMDhms counting scheme free of midday discontinuities to achieve reverse compatibility with the existing SMPTE time code systems. This first requires a general solution to the local timekeeping problem.

Common calendar takes the form of defining a set of local timescales together with technical specifications for interoperable time stamps and associated calculation formulas. By carefully defining the environment and parameters, it is possible to create a self-contained environment that is both reverse compatible to existing timekeeping technologies and provides a set of deterministic local timescales. This includes counting modes that can eliminate midday YMDhms discontinuities, as described in the section "TOD Count Mode: Leap Second at Midnight", which meets the initial goal of supporting legacy SMPTE time code requirements.

The integrity of the design has been verified in a prototype c++ reference application.

There are, of course, many design decisions implicit in the common calendar definitions and specifications that rely on interpretation of existing timekeeping standards and conventions. It is here common calendar may be improved with input from the wider timekeeping community, and the author looks forward to reader's thoughts and comments.

Conclusion

There are several issues with traditional timekeeping that prevent accurate time stamps in all cases.

Traditional timekeeping systems:

1. Operate on 86,400 second calendar days, while UTC may have 86,401 or 86,399 second days. This discrepancy underlies all difficulties with traditional systems. It cannot be resolved within the context of the traditional time format because there is no way to represent a positive leap second. This results in at least one indeterminate time stamp value at or near the leap second.
2. Complex and arguably vague specifications of UTC and mismatched or incomplete specifications of traditional systems contribute to misunderstanding and implementation errors and omissions.
3. Specifications of local time are not well defined, and widely deployed systems use different schemes (e.g., Windows vs. tz database).
4. The common practice of introducing leap seconds in the YMDhms encoding of local time simultaneous with introduction at UTC generates a discontinuity in midday of the local YMDhms sequence. This creates a complex asymmetrical matrix of local YMDhms encodings.
5. No binary time stamp format has the necessary and sufficient information to fully and deterministically describe local time.
6. YMDhms character encoding formats, often relying on ISO 8601 guidelines, have incomplete information to deterministically represent local time.

These factors combine to a witches' brew of complexity and controversy with resulting inaccuracies among systems.

Despite this, none of these conventions or practices can be changed without the risk of upsetting vast numbers of devices and applications.

1. It seems unlikely that ITU-R or any other organization can substantially change UTC, such as "eliminate leap seconds." This leaves the mismatch between 86,400 second day systems and UTC in place for the foreseeable future, certainly a decade.

2. The systems and applications using the 86,400 second day cannot not be suddenly replaced. The huge scale of deployment of these systems makes any near-term upgrades entirely infeasible.
3. The practice of introducing leap seconds in midday of local time simultaneous with UTC cannot be withdrawn without upsetting the many systems designed to expect it.
4. Definitions of local time will remain subject to change by local custom or political influence.
5. The practice of "leap second smear" used to avoid possible system disruption caused by the mismatch between 86,400 second day systems and UTC will probably continue and expand, further compromising time stamp accuracy.

If a worldwide accurate and deterministic timekeeping system is to emerge, it will only be realized through a long process of adoption and migration to some proposed and widely accepted plan. That plan must include an underlying design capable of supporting legacy timekeeping systems while providing for the complete and deterministic representation of local time that accounts for both UTC leap seconds and local time UTC offset and Daylight Saving rules. Any plan must be backed with rigorous technical specifications, attract wide adoption, and ultimately be standardized by international standard bodies.

The author hopes common calendar may provide a starting point for discussion, refinement, and formal standardization that may find its way to international acceptance.

Acknowledgments The Science of Time Symposium provided an informative and stimulating event for the presentation of common calendar that further informed the work. The author wants to thank all the participants for their comments and inspiration.

Thanks to the many contributors on the LEAPSECS discussion list.

Thanks to Steve Summit for his comments on the paper and his work in porting the common calendar c/c++ program to Linux.

References

Date and Time Terms and Definitions http://edlmax.com/Date_and_Time_Terms_and_Definitions.pdf

Supplementary Materials for Common Calendar http://edlmax.com/Common_Calendar_Supplement.pdf

BIPM The International System of Units (SI) 8th edition 2006 (commonly called the SI Brochure) http://www.bipm.org/en/publications/si-brochure/

BIPM JCGM 200:2012, International vocabulary of metrology – Basic and general concepts and associated terms (VIM) http://www.bipm.org/en/publications/guides/vim.html

ISO 8601 2004-12-01, Data elements and interchange formats — Information interchange — Representation of dates and times http://www.iso.org/iso/iso8601

IEC INTERNATIONAL STANDARD, 60050-11, Amendment 1 International Electrotechnical Vocabulary – Part 111: Physics and chemistry http://www.mz3r.com/fa/wp-content/uploads/2012/02/books/standards_of_IEC/60050-111.pdf

IEC INTERNATIONAL STANDARD, 60050-11, Amendment 1 International Electrotechnical Vocabulary –PART 713: RADIOCOMMUNICATIONS: TRANSMITTERS, RECEIVERS, NETWORKS AND OPERATION http://www.iea.lth.se/internt/IEC_Dictionary/Base/713.pdf
ITU-R TF.460-6 (02/02), Standard-Frequency and Time-Signal Emissions http://www.itu.int/rec/R-REC-TF.460/en
Society of Motion Picture and Television Engineers (SMPTE) https://www.smpte.org/
SMPTE ST 12-1:2014, Time and Control Code http://ieeexplore.ieee.org/document/7291029/
Conversion between SMPTE hh:mm:ss:ff Time Code and Frames http://edlmax.com/SMPTETimeCodeConversion.htm
ST 2059-2:2015 - SMPTE Standard - SMPTE Profile for Use of IEEE-1588 Precision Time Protocol in Professional Broadcast Applications http://ieeexplore.ieee.org/document/7291608/
ST 2059-1:2015 - SMPTE Standard - Generation and Alignment of Interface Signals to the SMPTE Epoch http://ieeexplore.ieee.org/document/7291827/
1588-2008 - IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems http://ieeexplore.ieee.org/document/4579760/
LEAPSECS -- Leap Second Discussion List https://pairlist6.pair.net/mailman/listinfo/leapsecs
The Science of Time Symposium http://sot2016.cfa.harvard.edu/

Chapter 35
The Transfer of Earth-Time to the Planets

David E. Smith and Maria T. Zuber

Abstract Understanding the development and future of the solar system requires an appreciation of the processes and physical laws that govern both motion and time in a gravitational environment. Processes and changes that appear small on human timescales may have a significant impact in the long term, and understanding them requires measurements of extreme accuracy. The history of the evolution of the solar system is manifest in the present-day motions and positions of the planetary bodies and the forces that helped shape them. Position and time are interrelated, and it is impossible to provide meaning to a precise position unless the precise time is known. Here we describe recent space experiments that attempt to synchronize events that take place at other planets with time on Earth, how time and distance are interrelated, and how being able to transfer time accurately and measure the positions of the planets can help understand the future of the solar system.

Keywords Interplanetary ranging • Planetary perturbations • Ranging experiments • Transferring time • Solar GM

Introduction

The need for knowing Earth-time for events that occur at other solar system bodies may not be obvious but for certain science objectives. Interpretation of measurements that quantitatively describe phenomena in astronomy, geophysics, geodesy, and fundamental physics necessitates knowledge of time in a common reference system. Nearly all these disciplines have a requirement for knowledge of the position being observed and the position from which the observations are made in addition to the "time" of the observation at both locations. The ephemerides of solar system bodies are usually developed in a time system that enables the prediction of the locations of these bodies as viewed from Earth, the observer's time system, but when the observations are being obtained at another body, it is necessary to rely on

D.E. Smith (✉) • M.T. Zuber
Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
e-mail: smithde@mit.edu; david.e.smith@nasa.gov

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_35

a clock at the planetary body or an orbiting spacecraft that is running independently of clocks on Earth. Thus, understanding the circumstances of an observation requires that the time at the planetary body must be correlated as precisely as possible with time on Earth.

Clocks of any kind are not stable and drift over time, and so relating Earth-time to time on a spacecraft at Jupiter, for example, requires a method of comparing two clocks separated by several ua. When two clocks are side-by-side, it is simple to adjust the time reading of one to match the other, and this can be done regularly or as needed. But clocks that are several ua apart require accurate knowledge of the distance between them in order to compensate for the time it takes for a signal to reach its destination. At Jupiter, about 5 ua from Earth, the time delay is of order 40 min. But if we wish to compare clocks at the nanosecond level we need to know the distance to 30 cm. So comparing time at two different locations becomes a problem of knowing or measuring the distance between the two clocks, and this is independent of any relativistic corrections arising from the relative velocities of the two clocks and the different gravity fields in which they are immersed. With knowledge of the distance, the transit time for a signal to reach its intended destination is known, and it becomes possible to compare the two clocks and "transfer" Earth-time to another planet. Opportunities for measuring interplanetary distances were discussed in Smith et al. (1997, 2015).

How well we need to know Earth-time in the outer solar system depends on the science objective to be accomplished. Here we argue that 0.5 ns (15 cm) will make a significant advance for geophysics and probably relativistic objectives.

Measuring Distance on an Interplanetary Scale

To measure the distance between two moving bodies requires knowing the time of an event at both bodies on the same time reference, however, knowledge of the time requires knowledge of the distance between the clocks. The general form of this problem was addressed by Degnan (2002, 2006) who showed that if you could send signals in each direction and compare the times and distances measured, it was possible to derive both the time difference and the distance between the clocks. This idea is illustrated in Fig. 35.1. In addition, if it is possible to repeat the measurement, then it is possible to derive the line of sight velocity between the two clocks and/or the drift of one clock with respect to the other due to the aging of the oscillator.

Another method of deriving the distance is by using a transponder approach in which a signal from one end arrives at the other end and triggers a return signal which is timed back at the original end by the same clock. This closed-loop approach immediately provides the distance between the two clocks after correction for any delay in the transponder process and any drift in the clock over the round-trip time of flight.

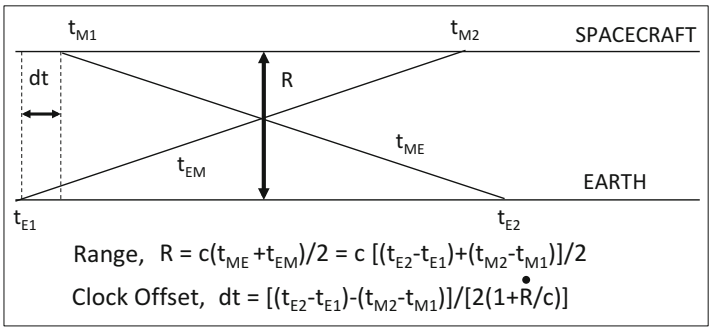


Fig. 35.1 Equations for an asynchronous transponder for deriving range and clock offset between two clocks. The parameter R is the distance between the clocks, c is the velocity of light, t_{ME} and t_{EM} are the times of flight of the uplink and downlink laser pulses, t_{E1} and t_{E2} are the times recorded at the Earth station, and t_{M1} and t_{M2} are the times recorded at the spacecraft (after Degnan 2006)

Many attempts to make decimeter measurements from laser ranging station on Earth to a planetary spacecraft were attempted in the 1990s (Zuber 2006), but they were all unsuccessful for a variety reasons. The first successful measurement was to the MESSENGER spacecraft during its cruise phase to the planet Mercury (Smith et al. 2006; Sun et al. 2013). The MESSENGER spacecraft was launched to Mercury in 2004, and it carried the Mercury Laser Altimeter (MLA) (Cavanaugh et al. 2007). In May 2005, the spacecraft performed an Earth gravity assist, and during its outward trajectory, a two-way laser ranging experiment was conducted between a laser tracking station at Greenbelt, Maryland, and the MLA instrument on MESSENGER. At the time of the measurement, MESSENGER was approximately 24 million kilometers from Earth, and the experiment was operated as an asynchronous transponder (Degnan et al. 2002) in which the Greenbelt laser station and the MLA on MESSENGER transmitted laser pulses independently toward each other. The experiment took place between May 27 and 31 and yielded two successful attempts at receiving laser pulses from MLA and one successful detection by MLA of pulses from Earth. The total number of pulses received was slightly more than 100, which was sufficient to calculate the range between the two lasers, their separation velocity and acceleration, and the offset between the clock on the ground and the clock on the spacecraft to less than a nanosecond. Table 35.1 summarizes the results.

The limitation of the results of the Earth to MLA experiment was directly proportional to the accuracy of the measurements made at the spacecraft, which were of order 0.5×10^{-9} s, which placed a limit on the range accuracy to be no better than 15 cm. Had the MLA instrument been of higher accuracy or precision then it may have been possible to achieve an even more accurate distance and time correlation of the Earth and spacecraft clocks and their relative drift rate.

A second experiment was conducted in September 2005 between the Earth-based laser station at Greenbelt, Maryland and the Mars Orbiter Laser Altimeter (MOLA) (Zuber et al. 1992) on the Mars Global Surveyor (MGS) spacecraft in orbit

Table 35.1 Result of the asynchronous transponder experiment between the Greenbelt laser ranging station and the Mercury Laser Altimeter (MLA) on the MESSENGER spacecraft in May 2005 (Smith et al. 2006)

Parameter	Laser solution	Uncertainty
Range	23,964,675,433.9 m	±0.2 m
Range rate	4154.663 m s^{-1}	±0.144 m s^{-1}
Acceleration	−0.0102 mm s^{-2}	±0.0004 mm s^{-2}
Time	71,163.729670967 s	$\pm 6.6 \times 10^{-10}$ s
Spacecraft clock drift rate	1.00000001559 ppb	$\pm 4.8 \times 10^{-10}$ ppb

at Mars, a distance of 80.1 million kilometers (Abshire et al. 2006). At this time, the MOLA instrument was no longer active as an altimeter as a result of a failure of the oscillator that commanded the laser to fire, but MOLA was active in its radiometer mode and was able to receive laser pulses from Earth. On September 28, 2005 the MOLA instrument successfully detected >500 laser pulses from Earth. Since there was no laser transmitting to Earth, it was impossible to accomplish a transponder experiment and compute accurately the distance, but analysis of the received pulse at MOLA did permit an approximate estimation of the difference between the spacecraft clock that was time tagging the receive pulses and Earth-time to between 120 and 140 μs, compared to less than 1 ns achieved in the two-way experiment to MLA a few months earlier.

The primary accomplishment of this experiment was the demonstration of the 80 Mkm link to Mars. The Mercury and Mars laser link experiments also represented achievements for the respective spacecraft. Because neither MESSENGER nor MGS had the requisite pointing accuracy to point and stare at a fixed location with the accuracy that the 100 μrad laser beams required, both needed to scan the laser telescope across the Earth in the general location of the Earth laser station. Figure 35.2 shows the scanning pattern of the MGS spacecraft at Mars to ensure that the MOLA telescope would be pointing at the laser station for some of time during the experiment.

In addition, neither the MLA nor the MOLA instrument was designed for the two-way observations that were eventually undertaken. Both experiments took advantage of the presence of laser systems at or on their way to other planetary bodies that were designed to make precision altimetry measurements at their destinations.

In June 2009, the Lunar Reconnaissance Orbiter (LRO) mission was launched to the Moon. It carried the Lunar Orbiter Laser altimeter (LOLA) (Smith et al. 2010) and a laser ranging system (LR) (Zuber et al. 2010) that worked in conjunction with the altimeter. Because of concerns that the orbit determination from routine Doppler tracking might not be adequate for the laser altimeter, the LR system was included so that laser ranging from Earth could improve the positioning of the LRO spacecraft. LR was a one-way ranging system that would operate at 28 Hz and pulses received by the LR system would be timed by LOLA (Fig. 35.3).

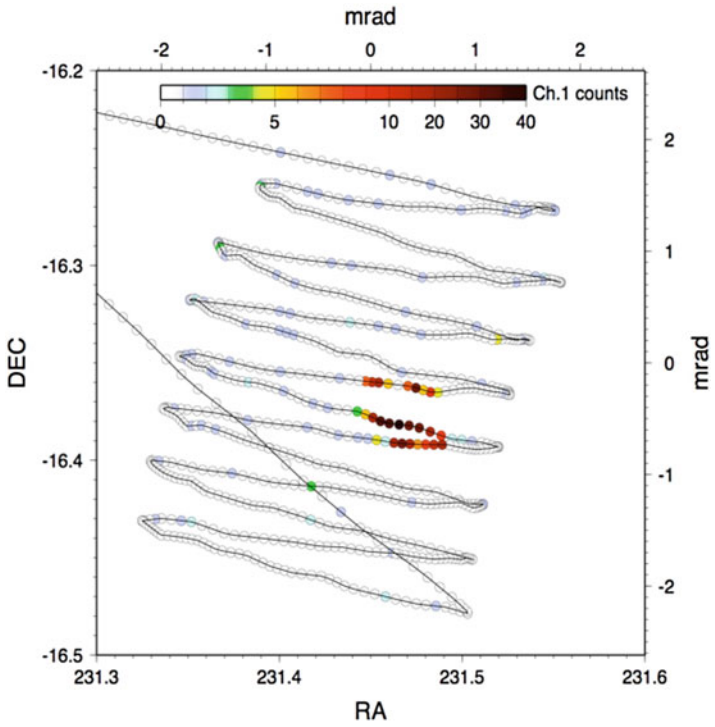


Fig. 35.2 The scan pattern performed by the MGS spacecraft of Earth centered on the laser tracking station at Greenbelt, Maryland on Sept 28, 2005. The successful detection of laser pulses from Earth is shown as colored spots per 125 ms. Over 500 laser pulses from Earth were detected by the spacecraft

The nominal accuracy of the system, which was limited by LOLA timing system, was 10 cm, but because it was a one-way system, the measurement was a biased range with a precision of 10 cm. However, because the spacecraft was in orbit about the Moon, whose distance was well known, the range bias was easily estimated from the data (Mao et al. 2016). Further, the regular estimation of the range bias could be used to derive the drift in the spacecraft clock (Fig. 35.4) used in timing the laser pulses from Earth.

The laser tracking of LRO by an international group of laser tracking stations continued routinely for 5 years (Mao et al. 2016) and besides fulfilling its intention of providing additional tracking and improving orbit determination of LRO (Bauer et al. 2016), it also enabled all LRO observations to be analyzed on Earth-time. Although less important due to the proximity of the Moon to the Earth, it demonstrated that transferring time from Earth to another solar system body could be accomplished routinely and with an accuracy of a nanosecond.

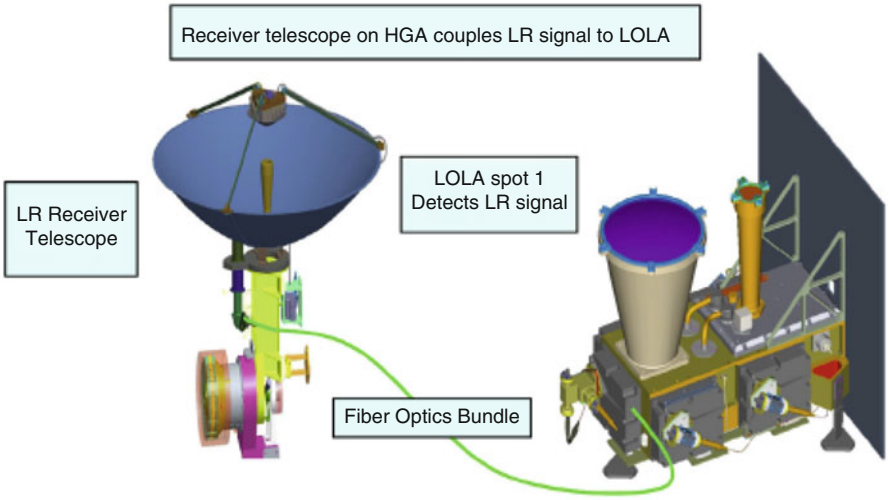


Fig. 35.3 The LR telescope was part of the high-gain antenna. Photons received by the LR telescope were transmitted to LOLA via a fiber optic cable where the signal was time-stamped

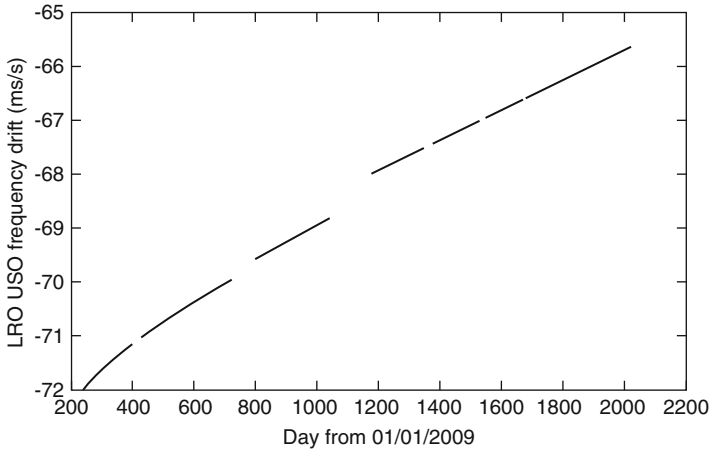


Fig. 35.4 The drift of the LRO oscillator over a 5-year period derived from the laser tracking of LRO enabled the time on the spacecraft to be related to time on Earth to about a nanosecond

Transferring Earth-Time to the Outer Solar System

The greater distance of the outer planets, such as Jupiter, presents additional technological challenges due, in part, to the time for a signal to reach its intended destination but does not change the requirement or the methodology by which it is possible to transfer time. For some objectives, a one-way system that is able to provide ms level timing is sufficient. In this case, the timing which is relative to the

velocity of the orbiting spacecraft, which typically may be a few km/s, will introduce errors of a few meters. But for other purposes in which we are interested in the change in position of a planet, it will require a transponder-type measurement of nanosecond accuracy. Relativistic studies may well require accuracies 10–100 times greater and present additional difficulties.

For understanding fundamental problems in Newtonian space, such as changes in GM of the sun, the product of the solar mass, and the gravitational constant, time and distance measurements of ns and cm are probably adequate, at least for the present. Possible changes in G have been considered for several decades (Van Flandern 1981; Anderson et al. 2015) with estimates of order $\delta G/G \sim 10^{-12}$ per Earth-year, although there is little or no observational evidence. A change in the mass of the sun would also affect the evolution of the solar system and would be indistinguishable from a change in G. With the estimated remaining lifetime of the sun about of 5×10^9 years, fractional changes of $<10^{-10}$ and smaller cannot be ruled out. Estimates based upon the conversion of hydrogen to helium suggest changes of order 10^{-14} per year, much smaller than the only estimate for a change in G.

However, a change in the solar GM due to a changing solar mass or a time dependence of G cannot be ruled out and could be constrained with accurate measurements of positions of planets. Even though it would not be able to tell which of the "constants" is actually changing. Changes at the magnitude of 10^{-12} year^{-1} could have significant consequences for our understanding of the origin and evolution of the solar system.

No matter what the cause, a change in the solar GM would cause the geometrical relationships between the planets to change with time, in addition to the changes due to their normal orbital relationships. Figures 35.5, 35.6, and 35.7 show how the distances from Earth of Mercury, Mars, and Jupiter change of over the period of a few years as a result of 1 part in 10^{12} year^{-1} change in the solar GM.

The essence of Figs. 35.5, 35.6, and 35.7 is that the changes in planetary distances from Earth are measureable with present technology as demonstrated by the experiments to Mercury, Mars, and the Moon, and the measurement of time at the planet is critical to the interpretation of the distance to the planet. Distance changes of decimeters to a meter over a few years are arguably measureable today. Further, since a change in the solar GM has the greatest effect on the inner planets, as we can see in Figs. 35.5–35.7, then a series of measurements from an outer planet to and inner planet is likely to show the largest change in the shortest time.

The distance from Earth of a terrestrial planet is probably most easily achieved by a lander on the surface, but for gaseous planets, such as Jupiter, the only likely way to make the necessary measurements is to a spacecraft in orbit about the planet. From the knowledge of the spacecraft orbit to the center of mass of the planet will require accurate positioning of the orbiting spacecraft and the time of the measurements, as was demonstrated by the laser ranging experiment on LRO at the Moon (Mao et al. 2016). Future missions to Jupiter and Mercury by ESA, BepiColombo to Mercury, and JUICE to Ganymede, both carry laser altimeters, and might be able to make these orbital measurements, although in neither case are measurements of the

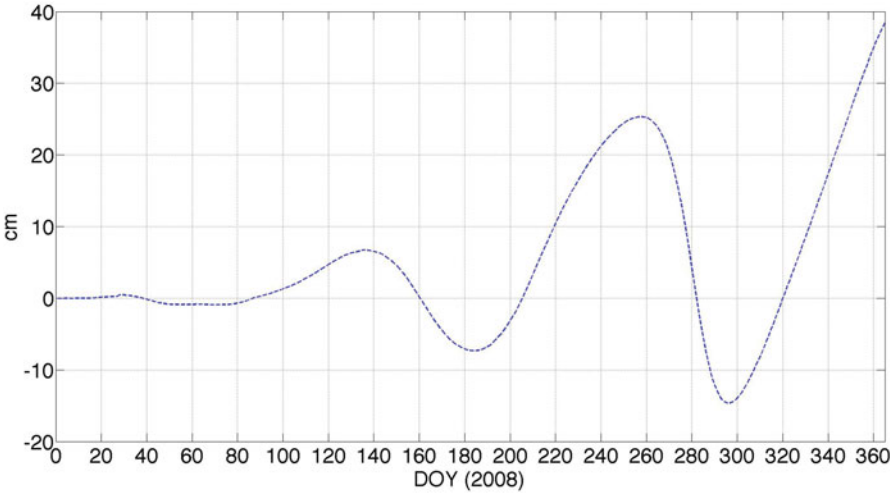


Fig. 35.5 The effect of $\delta G/G = 10^{-12}$ per Earth-year on the distance of Earth to Mercury in 2008. The oscillation is the result of Mercury's shorter orbital period around the sun, while the trend is a steady increase in separation of the two planets

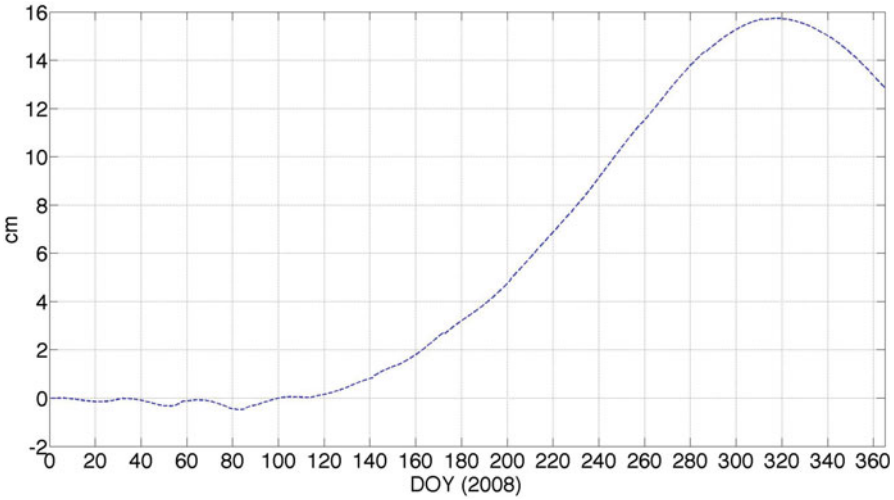


Fig. 35.6 The effect of $\delta G/G = 10^{-12}$ per Earth-year on the distance of Earth to Mars in 2008. The trend is to increase the distance between the planets modulated by the shorter period of Earth's orbital period around the sun

distance to Earth part of the present mission plan. NASA is also developing the Clipper mission to Europa and considering the possibility of a laser altimeter on a carrier relay spacecraft. Even if plans to make precise ranging and timing measurements are not part of the mission objectives, these missions do have the possibility

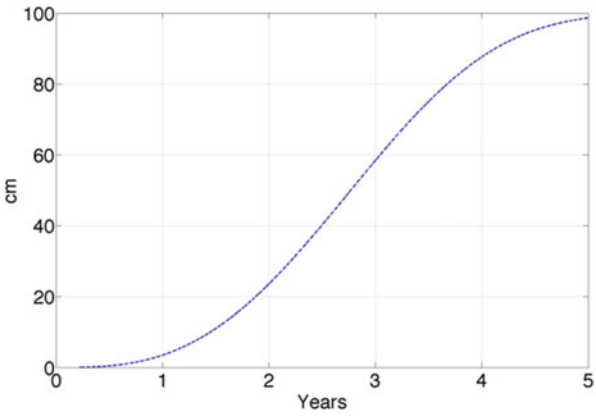


Fig. 35.7 The effect of $\delta G/G = 10^{-12}$ per Earth-year on the distance of Earth to Jupiter from 2008 to 2013. The distance between Earth and Jupiter increases monotonically over this 5-year period

of demonstrating the ability of optical ranging over distances of over 5 ua and would be important technical achievements.

Establishing that the scale of the solar system is, or is not, changing, or establishing an upper bound, would be a significant accomplishment and raise the important question of whether it is due to a change in mass of the sun, a change in gravitational constant, or possibly some other cause. Whatever the outcome, the result would be a major advance in our understanding of solar system dynamics.

Summary

In the last few years, it has been demonstrated that it is possible to make accurate range measurements, and transfer time, over planetary distances. Experience so far has indicated that measurements of time and range at the accuracies needed, over several ua, will be challenging and will be required to span several Earth-years. Further, it is known that planetary orbits are continually being disturbed at the accuracies described herein by other objects, and the separation of the various dynamical processes will be a major but exciting task for the science community.

Acknowledgments We are pleased to acknowledge Dr. Antonio Genova (MIT), Dr. Gregory Neumann (GSFC), and Dr. Erwan Mazarico (GSFC) for their assistance in the preparation of this paper.

References

J.B. Abshire, X. Sun, G. Neumann, J. McGarry, T. Zagwodski, P. Jester, H. Riris, M. Zuber, D. Smith, Laser pulses from Earth detected at Mars, in *Conference on Laser Electro-Optics (CLEO)*, Long Beach, CA, May 2006 (2016). doi:10.1109/CLEO.2006.4628090

J.D. Anderson, G. Schubert, V. Trimble, M.T. Feldman, Measurements of Newton's gravitational constant and the length of day. EPL **110**(2015), 10002 (2016). doi:10.1209/0295-5075/110/10002

S. Bauer, J. Oberst, D. Dirkx, M.H. Torrence, J.F. McGarry, G.A. Neumann, D. Mao, D.E. Smith, E. Mazarico, M.T. Zuber, H. Hussmann, Demonstration of orbit determination for the Lunar Reconnaissance Orbiter using one-way laser ranging data. Planet. Space Sci. **129**, 32–46 (2016). doi:10.1016/j.pss.2016.06.005

J.F. Cavanaugh, J.C. Smith, X. Sun, A.E. Bartels, L. Ramos-Izquierdo, D.J. Krebs, J.F. McGarry, R. Trunzo, A.M. Novo-Gradac, J.L. Britt, J. Karsh, R.B. Katz, A. Lukemire, R. Szymkiewicz, D.L. Berry, J.P. Swinski, G.A. Neumann, M.T. Zuber, D.E. Smith, The Mercury Laser Altimeter instrument for the MESSENGER mission. Space Sci. Rev. **131**, 451–480 (2007). doi:10.1007/s11214-007-9273-4.

J.J. Degnan, Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements. J. Geodyn. **34**, 503–549 (2002)

J.J. Degnan, Asynchronous laser transponders: a new tool for improved fundamental physics experiments, in *Quantum to Cosmos: Fundamental Physics Research in Space* (2006), pp. 265–278. doi:10.1142/9789814261210_0021

D. Mao, J.F. McGarry, E. Mazarico, G.A. Neumann, X. Sun, M.H. Torrence, T.W. Zagwodski, D.D. Rowlands, E.D. Hoffman, J.E. Golder, M.K. Barker, D.E. Smith, M.T. Zuber, The laser ranging experiment of the Lunar Reconnaissance Orbiter: five years of operations and data analysis. Icarus **283**, 55–69 (2016). doi:10.1016/j.icarus.2016.07.003

D.E. Smith, M.T. Zuber, J.J. Degnan, J.B. Abshire, Prospects for centimeter laser ranging throughout the solar system. Presented at the Eighth Marcel Grossmann Meeting on General Relativity, Jerusalem, 22–27 June 1997

D.E. Smith, M.T. Zuber, X. Sun, G.A. Neumann, J.F. Cavanaugh, J.F. McGarry, T.W. Zagwodzki, Two-way laser link over interplanetary distance. Science **311**, 53 (2006). doi:10.1126/science.1120091, 2006.

D.E. Smith, M.T. Zuber, G.B. Jackson, H. Riris, G.A. Neumann, X. Sun, J.F. Mcgarry, J.F. Cavanaugh, L.A. Ramos-Izquierdo, R. Zellar, M.H. Torrence, E. Mazarico, J. Connelly, A. Matuszeski, M. Ott, D.D. Rowlands, T. Zagwodzki, M.H. Torrence, R. Katz, I. Kleyner, C. Peters, P. Liiva, C. Coltharp, S. Schmidt, L. Ramsey, V.S. Scott, G. Unger, D.C. Krebs, A.-M. Novo-Gradac, G.B. Shaw, A.W. Wu, The Lunar Orbiter Laser Altimeter investigation on the Lunar Reconnaissance Orbiter mission. Space Sci. Rev. **150**, 209–241 (2010). doi:10.1007/s11214-009-9512-y

D.E. Smith, M.T. Zuber, G.A. Neumann, E. Mazarico, X. Sun, Science and high accuracy laser ranging between the planets. Presented EGU General Assembly 2015, EGU-7736

X. Sun, D.R. Skillman, E.D. Hoffman, D. Mao, J.F. McGarry, L. McIntire, R.S. Zellar, F.M. Davidson, W.H. Fong, M.A. Krainak, G.A. Neumann, M.T. Zuber, D.E. Smith, Free space laser communication experiments from Earth to the Lunar Reconnaissance Orbiter in lunar orbit. Opt. Express (2013)

T.C. Van Flandern, Is the gravitational constant changing? Astrophys. J. **248**, 813–816 (1981)

M.T. Zuber, Seconds of data, years of trying. Photon. Spectra, 56–62 (May 2006)

M.T. Zuber, D.E. Smith, S. Soloman, D. Muhleman, J. Head, J. Garvin, J. Abshire, J. Bufton, Mars observer laser altimeter investigation. J. Geophys. Res. **97**(E5), 7781–7797 (1992)

M.T. Zuber, D.E. Smith, R. Zellar, G.A. Neumann, X. Sun, J. Connelly, A. Matuszeski, J.F. McGarry, M. Ott, L. Ramos-Izquierdo, D.D. Rowlands, T. Zagwodzki, M.H. Torrence, The Lunar Reconnaissance Orbiter laser ranging investigation. Space Sci. Rev. **150**, 63–80 (2010). doi:10.1007/s11214-009-9511-z

Chapter 36
Keeping Time with the Asteroids

Rob Seaman, Frank Shelly, Eric Christensen, Alexander Gibbs, and Stephen Larson

Abstract The Catalina Sky Survey (CSS) has discovered 5790 of 13,392 known near-Earth asteroids (as of writing), as well as tens of thousands from the main belt, and over 300 comets. Astrometric observations from the three CSS telescopes have contributed to computing the orbits of the great majority of the 700,000 known asteroids of all types. CSS remains the only NEO survey to have discovered asteroids prior to Earth impact.

The distinctive feature of CSS operations is near real-time reporting and follow-up of new discoveries. Minutes after an NEO is first detected, astrometry is submitted to the Minor Planet Center for posting on the NEO Confirmation Page providing targeting coordinates for observers around the world. Rapidly moving NEOs risk becoming lost again unless observed over a sufficient orbital arc; a good orbit solution is required for future study of any solar system object.

The engineering trade-offs are thus between tuning the observing parameters maximizing discoveries of new objects, optimizing time-tagged astrometric and photometric precision, and a workflow supporting both survey and follow-up observations. We will discuss an automated tool, AsteroidClock, which compares statistics for known asteroids in the field of view against measurements taken through the survey workflow. The ephemerides of previously discovered objects serve as ground truth to evaluate the many degrees of freedom inherent in tuning the survey's discovery and astrometry algorithms. As the name implies, an inventory of asteroids in the field of view can also provide a unique time stamp.

Keywords Asteroids • Ephemerides • Orbits • Astrometry • Orrery

The Catalina Sky Survey (CSS) has discovered 5790 of 13,392 known near-Earth asteroids (as of writing), as well as tens of thousands from the main belt, and over

R. Seaman (✉) • F. Shelly • E. Christensen • A. Gibbs • S. Larson
Lunar and Planetary Laboratory, International Bureau for Weights and Measures, University of Arizona, Arizona, USA
e-mail: seaman@lpl.arizona.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_36

300 comets. Astrometric observations from the three CSS telescopes have contributed to computing the orbits of the great majority of the 700,000 known asteroids of all types. CSS remains the only NEO survey to have discovered asteroids prior to Earth impact.

Chapter 37
Long-Term Timekeeping in the Clock of the Long Now

W. Daniel Hillis

Abstract We construct our systems of time for many different purposes: to measure intervals and events, to predict the future and solidify the past, and to align our activities with the cycles of the seasons and days. For most purposes, we can ignore the differences between the kinds of time that we need for these different purposes, but over long intervals, we cannot. This article describes the nine kinds of time that are used in the construction of the Clock of the Long Now, a mechanical clock that is designed to keep time continuously for 10,000 years.

Keywords Equation of time • Clock of the long now • Ten-thousand-year clock • 10,000-year clock • Mechanical timekeeping • Pendulum clocks • Solar sychronization

The purpose, design, and construction of the Clock of the Long Now is described elsewhere (Hillis 1995; Hillis 2013; http://www.longnow.org; Brand 1999; Hillis 2000). Besides being designed for an operation lifetime of 10,000 years, the clock has several other unusual features. Like many clocks, it displays the positions of the Sun, Moon, and stars, as well as the time of day and Gregorian calendar date. Unlike most clocks, it only updates these displays when the clock is wound. When the clock is visited (it is buried deep inside a mountain), it initially shows the date and time of the last visitor. Only when it is wound does the display advance to the current date and time. The clock must keep track of passing time even when the display is not visited. The small power required to keep the pendulum swinging is generated by temperature changes from night to day. Only the power required to advance the display is generated by human winding. This allows the visitor to watch the display being updated at an accelerated rate.

The clock has a set of chimes which ring in a different permutation each day, so that the order of sequence never repeats over the lifetime of the clock. The chime mechanism only operates at solar noon, and only if the clock is fully wound. This can occur either because someone has just recently wound the clock or because the

W.D. Hillis (✉)
The Long Now Foundation, San Francisco, CA, USA
e-mail: alice@appliedinvention.com

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_37

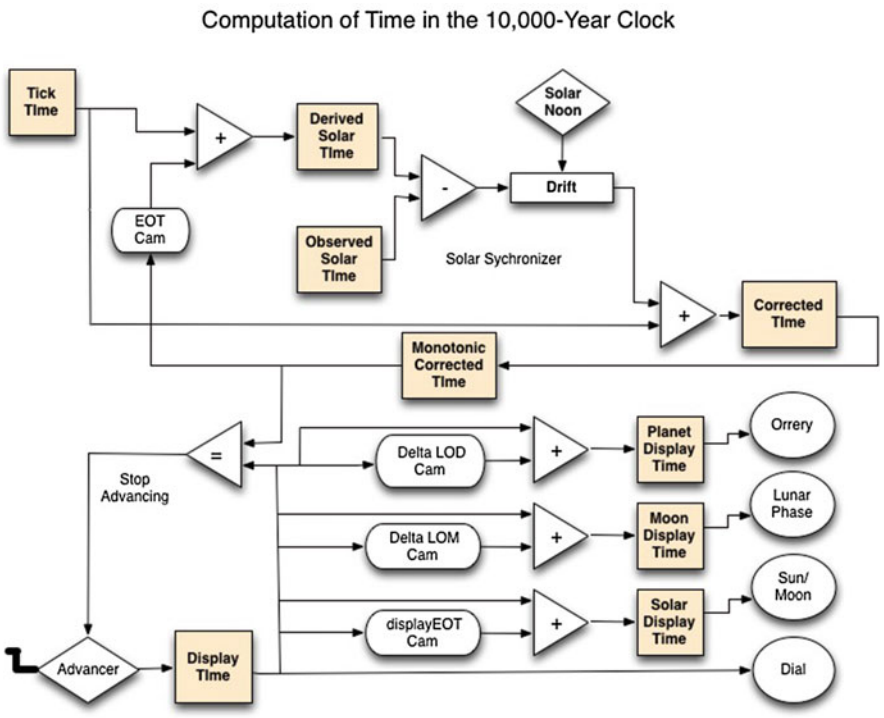


Fig. 37.1 Computation of time in the 10,000-year clock

self-winding mechanism that is powered by day-night temperature swings has accumulated extra power.

A second unusual feature of the clock is that it is phase synchronized to the Sun. Every summer solstice, light shines down the shaft at solar noon, so that a lens focuses the light on a trigger that detects thermal expansion. This solar trigger is used to adjust the time that is generated from the swinging pendulum, correcting any long-term drift relative to the timing of the Sun.

These functions require that the clock keep track of both solar and dynamical time,[1] as well as the interval between the displayed time and the current time,[2] and other rates required for operating the astronomical displays and the chimes. In all, there are nine kinds of time that are represented in the Clock of the Long Now. The relationship between them is show in Fig. 37.1 and explained below.

Tick time is the time generated by counting the swings of the pendulum, which has a period of 7.5 s. This is used to derive the apparent solar time by adding a correction for the equation of time, that is, the difference between apparent solar and mean solar time, which is generated by a cam. This cam is different from the

[1]That is, time responsive to the slowing Earth versus the steady cadence of the pendulum.

[2]The precision of the clock is 5 min (in whatever time scale) allowing us a certain latitude of expression in this paper.

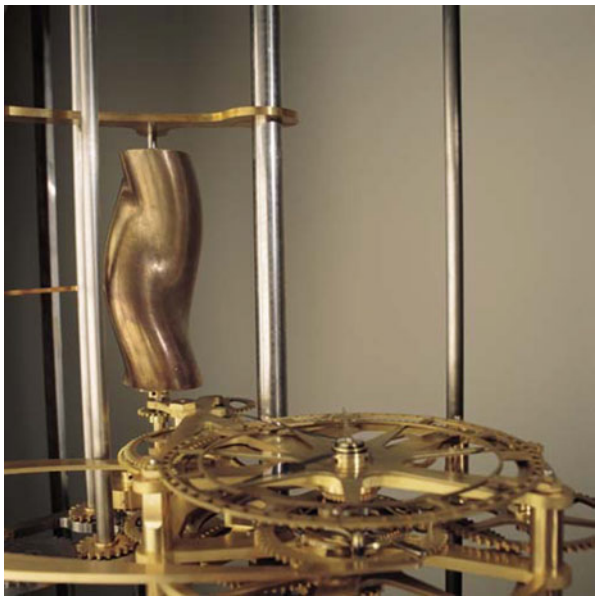


Fig. 37.2 An early prototype of the equation of time cam

equation of time cam in other clocks, because it must take into account the way that the equation of time changes slowly over the course of millennia, due to the factors such as precession of the Earth's axis, the slowing of its rotation, and the precession of the Earth's orbit around the Sun. A prototype cam is shown in Fig. 37.2. The cam rotates roughly once a year, and the cam follower rises along the length of the cam over the course of 10,000 years. Computing the shape of this cam required extending the JPL Ephemerides (Folkner et al. 1994; Hillis et al. 2011), 10,000 years into the future.

The correction from the equation of time cam is added to tick time to generate *derived solar time*. Since derived solar time is used to rotate the cam, this creates a circularity in its definition. This is not a problem since the correction to the "gain" around the control loop is small.

Observed solar time is based on the observation of the Sun. It is only valid on clear days at solar noon near the time of the summer solstice. This raises the question of how often such conditions may fail to occur, given the possibility of weather patterns changing substantially over the design lifetime of the clock. Based on the predicted accuracy of the pendulum, the clock may require a synchronization event as often as once every 22 years. This is not currently a danger in West Texas, where the clock is located, but the weather patterns may change. An investigation was performed to see if a shift in weather patterns was likely to cause such a gap (Grossman et al. 2014). This was accomplished by using historical weather data from all over the globe and asking if there is any place in the current world where the clock would have a significant chance of failing to synchronize often enough. For example, Fig. 37.3 shows locations where an inferior clock would not work, if it required a synchronization event every 8 years. The historical data showed no place on Earth where current weather patterns would produce more than a 1% probability

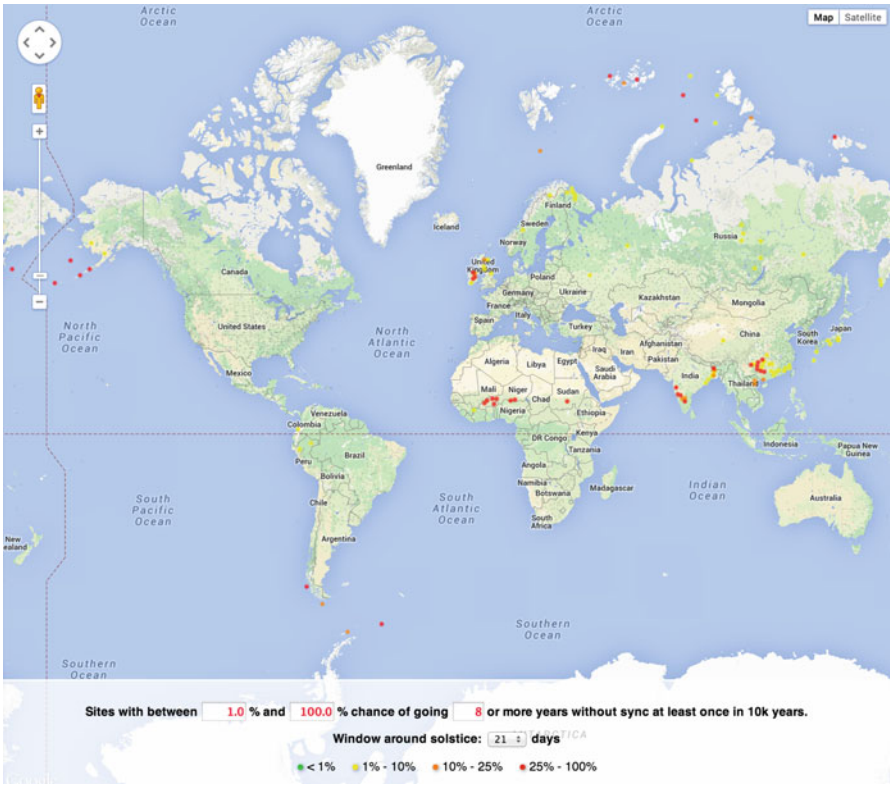


Fig. 37.3 Locations where synchronization of a hypothetical inferior clock would fail because of 8 consecutive years of blocked summer solstice Sun with >1% probability, sometime in 10,000 years

of 22 consecutive years of blocked summer solstice Sun in 10,000 years. Our conclusion is that shifting weather problems are not likely to cause a problem.

Whenever observed solar time is valid (the Sun is observed), the difference between it and derived solar time is used to correct the drift of tick time to produce *corrected time*. If the pendulum was perfectly accurate, corrected time would be the same as tick time. In reality, it is likely to occasionally jump forward or backward. Thus, corrected time is not necessarily monotonic and can repeat.

The non-monotonicity would be a problem for triggering events such as the chimes, since it might cause them to be re-triggered. For this reason, *monotonic corrected time* is generated from corrected time. Monotonic corrected time tracks corrected time, but pauses whenever corrected time reverses, until it catches up.

Display time is used to generate the time shown on the dial as well as the displays of the Sun, Moon and stars and the Gregorian calendar date. This time is only accurate immediately after the clock has been wound. It is updated by keeping track of the interval since the last winding, as measured by monotonic corrected time.

Table 37.1 Time periods displayed by the clock

	Julian days	Solar days	Clock	Time	Gear ratio
Sun/moon position	1.000000984	1	1	Solar	144 288
					12 12
Moon phase	0.966136786	0.966136786	0.966136789	Lunar	220 263
					19 197 16
Sidereal day	0.997270651	0.997269669	0.997269669	Solar	235 342
					220 263
					343 227 19
					197 16
Precession	9412982.24	9412972.980	9412882.619	Planet	329 282
					225 220
					317 121
					9 9 10 19
					19 64
Tropical year	365.242189	365.2415227	365.2415166	Planet	220 317
					121
					19 19 64
Sidereal year	365.2563554	365.2559961	365.2564103	Planet	56 111 220
					16 18 13
Mercury	87.96925718	87.96917062	87.96914701	Planet	321 151
					19 29
Calendar year	365.2428594	365.2425000	365.2425000	Planet	321 151
					235
Venus	224.7007992	224.7005781	224.7005862	Planet	19 29 92
					56111 220
Earth	365.2563554	365.2559960	365.2564103	Planet	16 18 13
					219 23 61
					121 97 111
					220
Mars	686.9796466	686.9789706	686.978544	Display	43 241 236
					14 16 18 13
					187 121 97
					111 220
Jupiter	4332.820129	4332.815866	4332.798497	Planet	236 14 16
					18 13
					238 97 111
					220
Saturn	10755.69864	10755.68806	10755.49679	Planet	14 16 18 13
					321 151

The "gear ratio" column shows the ratio of the gear train connecting the display to the indicated time base, approximating the number of the predicted solar days

The display for the position of the Sun requires the recomputation of apparent local solar time from displays through the addition of a correction generated by a second equation of time cam, similar to the first. This generates *solar display time*.

The display for the phase and position of the Moon requires a correction for the predicted changes in both the Earth's and the Moon's orbital rate. This generates the addition of a length of month correction cam. This generates *moon display time*.

The display for the planets' relative positions requires only a correction for the predicted changes in the Earth's rotation. This generates the addition of a length of day correction cam. This generates *planet display time*, which is the closest long-term approximation in the clock to Barycentric Dynamical Time. Without this correction, the computed position of Mercury would be significantly inaccurate in 10,000 years. The periods and the rational approximations by which they are calculated are shown in Table 37.1.

Thus, the Clock of the Long Now needs to generate nine versions of time to meet different requirements. Given that nine versions of time are required in a single clock, perhaps it is less surprising that so many different time standards are maintained by timekeepers throughout the world.

References

S. Brand, *The Clock of the Long Now: Time and Responsibility* (Basic Books, 1999)

P. Folkner, et al., Astron. Astrophys. **287**, 279–289 (1994)

A. Grossman, J. LaPorte, B. Walker, *Long Now Clock Weather Investigation*, May 2014, available from the Long Now Foundation

W.D. Hillis, *The Millennium Clock*, Wired, Scenarios: The Future of the Future, special edition, 1995

D. Hillis, *Introduction to the Long Now Clock Project*. Horological Science Newsletter, Feb. 2000, p. 2

D. Hillis, Compensating compound pendulums. Horol. J., 497–500 (2013)

D. Hillis, R. Seaman, S. Allen, J. Giorgini, in *Time in the 10,000-Year Clock*, ed. by J.H. Seago, R.L. Seaman, S.L. Allen, Decoupling Civil Timekeeping from Earth Rotation (American Astronautical Society, 2011)

Chapter 38
Aspects of Time Distribution

Martin Burnicki

Abstract A basic overview of how time can be synchronized in industrial and daily life applications, focusing on the types of reference clocks and time sources that can be used for time synchronization, as well as possible ways to distribute the time from these sources, and limitations of different approaches.

Keywords Time source • Accuracy • Precision • GPS • GLONASS • Galileo • Beidou • GNSS • DCF77 • WWVB • MSF • JJY • Long wave transmission • Signal propagation • Transmission delay • NTP • PTP • White Rabbit • Telephone modem services • ACTS • Spoofing • Jamming • PCI • USB • Serial connection • Latencies • Pulse-per-second • PPS • System time adjustment • Clock drift compensation • Leap second handling • Virtual machines • Time distribution • Time code • IRIG • IEEE 1344 • IEEE C37.118 • ntpd • loopstats • NTP Pool

Who Needs Time Synchronization?

Time synchronization is a strong requirement in many industrial applications as well as in daily life, even if this is often not obvious.

Figure 38.1 shows some typical applications which include:

1. Air traffic control	13. Lottery
2. Research vessels	14. Traffic management
3. Oil production	15. Operation coordination
4. Satellite communication	16. Event management
5. Observatories	17. Wall clocks
6. Power substations	18. Lighting control
7. Power plants	19. Railway timetable
8. Toll charging systems	20. Radio broadcasting
9. Wind energy plants	21. Mobile communication, call data records

(continued)

M. Burnicki (✉)
Meinberg Funkuhren GmbH & Co., KG, Bad Pyrmont, Germany
e-mail: martin.burnicki@meinberg.de

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_38

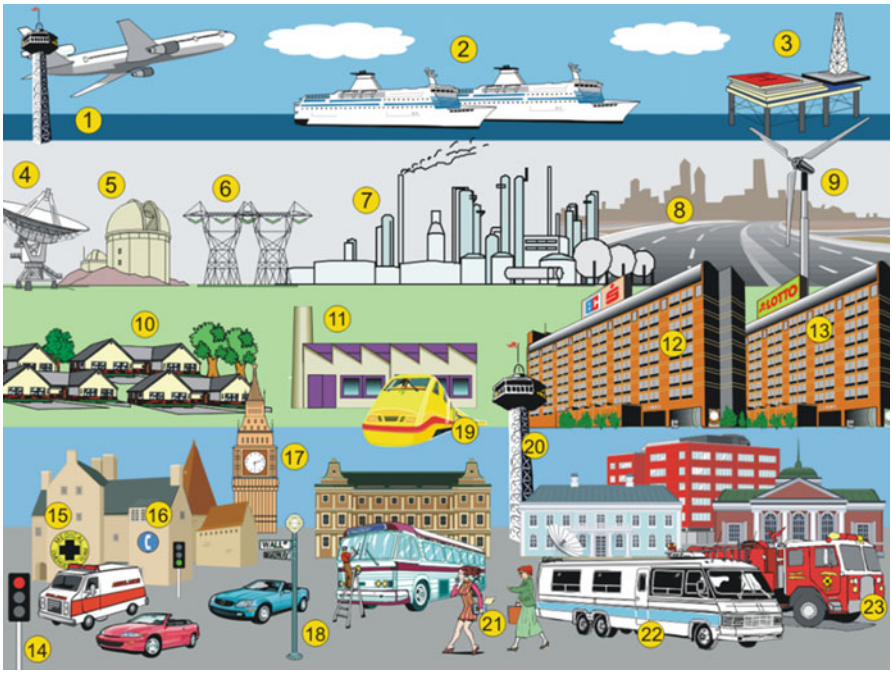


Fig. 38.1 Typical applications requiring time synchronization

10. Public infrastructure	22. Outside broadcast van
11. Production flow	23. Emergency
12. Banks, cash terminals, stock exchange, data centers	

Of course not every application requires the same level of accuracy, so different approaches can be used for different applications.

Basic Reflections When Planning Time Synchronization

Depending on the requirements of a particular application, there are a few basic questions:

- Which time sources are available, and which accuracy is provided by each time source?
- How to get the time from a time source into the system/application?
- Which accuracy can be achieved depending on the client characteristics?

These questions are discussed in the following chapters.

Reference Time Sources

There are a number of different time sources available, each of which has some specific characteristics:

- Long wave receivers (DCF77, WWVB, MSF, JJY, . . .)
- Telephone services (ACTS by NIST, PTB)
- GNSS satellite receivers (GPS, GLONASS, Galileo, Beidou, . . .)
- Time servers/network protocols (NTP, PTP, "White Rabbit")

Long Wave Receivers

A number of countries are operating long wave transmitters which broadcast the current time and eventually some additional information via a long wave signal. Examples are:

- *DCF77* in Germany
- *WWVB* in the United States
- *MSF* in the UK
- *JJY* in Japan

In many cases, the long wave signal can be received inside buildings, and receivers are cheap, so they can easily be integrated into wall clocks, alarm clocks, wrist watches, etc. However, the possible accuracy is usually limited to one or a few milliseconds, which is mainly due to the typical characteristics of long wave signals in general and the limited bandwidth available for transmission.

The limited bandwidth together with noise filtering in the receiver makes it hard to determine the point of time of a certain time mark very accurately. Figure 38.2 shows an example of a received long wave signal with narrow bandwidth (top) and broad bandwidth (bottom).

While a narrow bandwidth is better to suppress electrical noise, the slopes of the demodulated time marks become much flatter, and thus the beginning of a time mark can only be determined less accurately than with broad bandwidth and steeper slopes.

At least if accuracies better than 1 ms are required, then the propagation delay of the radio-frequency (RF) signal needs to be determined and compensated in the receiver. However, a long wave receiver is unable to determine the propagation delay by itself since usually only the signal from a single transmitter can be received. So the receiver can't find out how long the signal has been propagating before it arrived at the receiving location.

A basic problem with long wave signals is that there is a ground wave signal which propagates along the surface of the globe and a sky wave which is reflected at the ionosphere, as shown in Fig. 38.3. The ground wave signal is strong at distances up to a few 100 km around the transmitter, but at distances beyond this, the wave

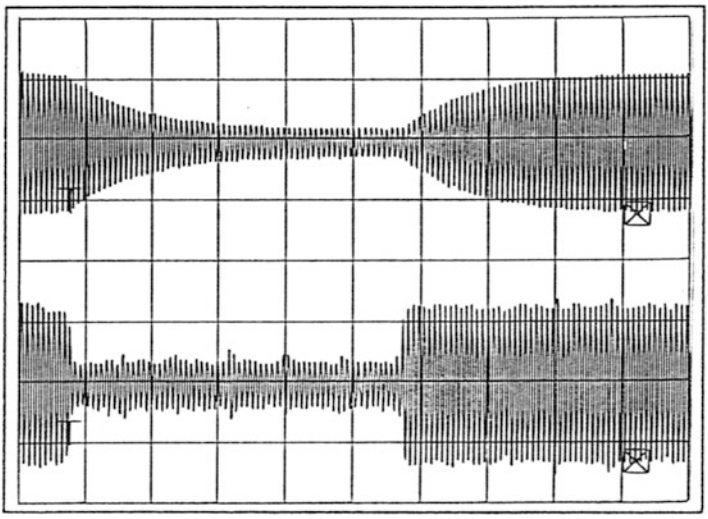


Fig. 38.2 Amplitude-modulated low-frequency signal with different bandwidths

Fig. 38.3 Low-frequency signal propagation

reflected at the ionosphere dominates, and obviously the way the reflected wave has taken is significantly longer than the direct way along the globe's surface.

So even if you try to compute the geographic distance of the receiver from the transmitter to compensate the propagation delay, the resulting accuracy depends on whether the ground wave or the sky wave is stronger at the location of the receiver.

Things get even worse if you take into account that the height of the ionosphere over ground varies with the season, the weather, and the time of day. This lets the real distance of the reflected long wave signal vary slowly over time. At distances from the transmitter where the ground wave and the sky wave arrive with a similar signal strength, the variation of the sky wave's signal path causes a phase shift of the sky wave signal relative to the ground wave signal, which in turn causes an interference between the two signals which in worst case leads to a cancelation of the combined signal whenever the phase shift is about 180°. As a result, reception

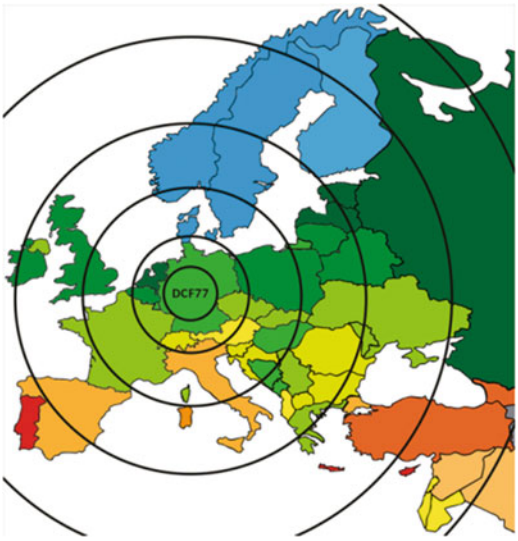


Fig. 38.4 Useable range of the DCF77 transmitter in Germany

may temporarily not be possible at all in the regions with the affected distance from the transmitter.

Long wave signals can only be received up to a certain distance from the transmitter, so another limitation is that these receivers can't be used worldwide. Figure 38.4 shows the useable range of the DCF77 transmitter in Germany.

An estimation of the WWVB's coverage is available at the NIST web site (NIST 2016).

Time signals provided by long wave transmitters use only very simple modulation and encoding schemes. These signals can easily be *jammed* or *spoofed* over relatively short distances at locations where the original signal from the transmitter is weak. *Jamming* means reception is disrupted, which can easily be detected, but *spoofing* is even worse since a wrong time can be supplied to receivers.

Telephone Modem Services

Some institutes provide the current time via a telephone modem service called *Automated Computer Time Service* (ACTS). This service is provided, for example, by the National Institute of Standards and Technology (NIST) in the United States and the Physikalisch-Technische Bundesanstalt (PTB) in Germany.

Whenever an ACTS client wants to adjust its time, it enables its modem and dials into the ACTS service. Since the dial-up connection has to be established first whenever a new query is made, it usually takes significant time until the first reply from the ACTS service is received, so clients usually send a few queries after each other via the modem line once the connection has been set up.

The advantage of the ACTS approach is that no antenna is required, just a telephone line. The drawback is that telephone costs may apply for each call to synchronize the time, and the achievable accuracy is usually in the range of one or a few milliseconds.

For example, the reference implementation of NTP, *ntpd*, can be configured to use an ACTS service as reference time source (NTP Documentation 2012).

Satellite Systems

Each of the *Global Navigation Satellite Systems* (GNSS) provides a whole set of satellites, which usually circle around the globe, so that always at least a couple of satellites of a particular system can be tracked from any particular location on Earth.

Popular GNSS systems are:

- *GPS*, the Global Positioning System operated by the United States
- *GLONASS*, operated by Russia
- *Galileo*, operated by the European GNSS Agency
- *BeiDou*, operated by China

Figure 38.5 shows the basic operation of a GNSS system.

As the name suggests, the main purpose of these GNSS systems is navigation, but the satellites also provide very accurate time, and the propagation delay of the RF signals can be compensated automatically by the GNSS receivers. These are preconditions for a receiver to be able to determine its own geographic position. At

Fig. 38.5 Basic operation of a GNSS system

least four satellites need to be tracked simultaneously to compute the receiver's position in space (x, y, z) and the current time (t).

There are current receiver models available which can even receive the signal from satellites of two or more different GNSS systems simultaneously, which provides a certain level of redundancy, or can be used for plausibility checks of the decoded time.

Commonly available GNSS receivers can yield an accuracy of a few nanoseconds, and there are special methods which allow for an even better accuracy.

Since the satellite signals use enhanced modulation techniques, they have a high level of noise immunity. However, these signals can also be jammed intentionally or unintentionally. Spoofing requires more effort than with long wave transmitters, but this is also possible.

NTP and PTP Time Services

NTP and PTP are network protocols where clients can get accurate time via a network connection. Individuals or companies can set up their own time servers, but there are also public services provided, e.g., by metrology institutes, and there is the NTP pool project where free NTP service is offered by volunteers (The NTP Pool Project 2016).

Servers providing time via NTP or PTP often have a built-in reference clock which gets the time from one of the sources mentioned earlier or even from another server accessible via the network.

Details of the network protocols are described in section "Network Time Synchronization Protocols".

Getting the Accurate Time into an Application

Once the accurate time is available from a particular time source, it needs to be accessed by or passed to an application. There are different ways to get the time from a radio clock or GNSS receiver, but there are also some advantages and limitations for the individual ways.

PCI Cards and USB Devices

With most PCI devices and some USB devices, it is possible to let an application read high-resolution time stamps from the device whenever required.

PCI cards can be addressed directly by the system processor, but USB devices need to be accessed via the operating system's USB support layer, so USB access is usually significantly slower than PCI access.

However, also PCI access is done via a peripheral system bus which may require some significant access time. This is described in detail in section "PCI Access Latencies".

Serial Connections/USB

Devices which are connected through a serial port or serial-to-USB converter usually send one or more different time strings in periodic intervals, but not necessarily at well-defined points in time. An application reading the time from such device has to:

- Wait until a time string comes in.
- Try to determine and compensate the transmission delay.
- Relate the time decoded from the string to a specific point in time.

The *transmission delay* to be compensated depends on the baud rate (e.g., 4800, 9600, or even 115200 baud) and framing (e.g., 8N1 or 8N2, resp. 10 or 11 bits transmitted per character) used for the serial connection but also on the UART chip and serial port driver. For example, UART chips commonly assembled in PCs have a built-in receive FIFO with a programmable threshold between 1 and 16 characters, so depending on the currently configured threshold, the driver and application may see a received character immediately after it has come in, or only after 16 characters have already arrived. So obviously the transmission delay to be compensated depends on the FIFO settings too.

If a serial-to-USB converter is involved, then the transmitted characters are put into USB messages which are forwarded by the computer's USB subsystem. This introduces additional delays which can't easily be determined and compensated, so the achievable accuracy is usually degraded.

The application also has to know (or make assumptions on) the timing relation of the string, e.g., if the string is always sent with a fixed offset after a second changeover, or at a random point in time during a second. A number of GNSS receiver modules send time strings in NMEA format without fixed timing relation.

If there's no fixed timing relation between the second changeover and the transmission of the time string, then a hardware signal, e.g., the slopes of a *pulse-per-second* (PPS) signal, can be used in addition to compensate the undefined timing relation of the string, if the *client hardware and software* support this.

Anyway, the behavior is similar to a church bell: when the bell rings or the time string arrives, the application knows which time it is *at this moment*. However, a few moments later it doesn't know the current time anymore unless it has its own internal clock which tells how much time has expired since the last event.

Many GNSS receivers provide a set of so-called NMEA strings, while other devices provide a proprietary data format from specific manufacturers. NMEA defines a number of different string formats, so-called *sentences*, which provide different sets of information. So care must be taken which sentences are transmitted and which of the transmitted sentences are to be used for time synchronization.

Pulse-Per-Second Signal

GNSS receivers usually output a *pulse-per-second* (PPS) signal which marks specific points in time, e.g., second changeovers. If the application's hardware and software support this, then the PPS signal can be used to compensate latencies from serial transmissions to increase the achievable accuracy.

An appropriate digital input must be available to evaluate a PPS signal. With standard PCs, a legacy printer port or the DCD pin of a serial port can be used as PPS input, if one of these ports is still provided by current PC hardware.

If a PPS signal is applied to such a port, then an interrupt request (IRQ) can be generated whenever a PPS slope is detected, and the interrupt service routine (ISR) can take a system time stamp associated with this slope. The time stamp can then be retrieved by an application, e.g., to compensate serial transmission delays.

There's still a delay from the moment where the IRQ is generated until it is serviced by the ISR, and the system time stamp is taken. This interrupt latency can be in the range of one to some tens of microseconds, depending on the host system, and causes a time offset error in the same range.

A possible way to avoid interrupt latencies would be to let the PPS slope latch the current count of a timer chip directly by hardware and then also evaluate this timer in the ISR. However, this can only be done with special hardware and isn't supported by standard PC printer or COM ports.

Linux, some BSD variants, and Solaris have built-in support for a PPS signal, which means a time stamp of the system time is taken by the kernel whenever a PPS slope is detected, and an application can retrieve the time stamp from the kernel. Thus, time synchronization software like the NTP daemon (ntpd) can use this signal to adjust the system time very accurately under these operating systems. On the other hand, for example, Windows has no built-in PPS support, and a special version of a serial port driver first needs to be installed to support this.

Most current computers don't provide a legacy printer or COM port anymore, so there are serial-to-USB converters available which provide a serial port via a USB connection. This usually works well for serial communication which is not related to accurate timing but doesn't help very much if a PPS signal is applied via such a converter. The reason is that the slope of a PPS signal applied to the DCD pin of such a converter only generates a USB message indicating that the level of the DCD pin has changed, and this is sent to the kernel via the normal USB subsystem, with all its latencies. Similar limitations apply to serial-over-LAN converters.

Characteristics of the Client System

The achievable accuracy on a client system does not only depend on the accuracy of the reference time source; it also depends strongly on a number of characteristics of the client hardware and software.

Resolution of the System Time

The resolution of the system time depends on the software as well as on the hardware of the target system.

For example, current Linux and other Unix-like systems provide microsecond or even nanosecond resolution. This means that different time stamps are returned by the kernel even if the system time is read in very short intervals.

On the other hand, on many Windows systems, the time resolution is limited to a timer tick, which is usually one or a few *milliseconds* only:

- Up to Windows XP/Server 2003: about 16 ms
- Since Windows Vista/Server 2008: 1 ms

For example, the *GetSystemTimeAsFileTime* function provided by Windows returns the current system time as *FILETIME* structure can hold time stamps with 100 ns resolution, but if an application calls this function repeatedly in a fast loop, then the *same time stamp* is returned during a whole timer tick interval.

Only since Windows 8/Server 2012, the Windows kernel keeps time with 100 ns resolution, but an alternate function *GetSystemTimePreciseAsFileTime* has to be used explicitly to retrieve precise time stamps. The legacy function *GetSystemTimeAsFileTime* has the same limitations as in older Windows versions.

It is important to keep in mind that nanosecond resolution does not necessarily mean nanosecond accuracy, but high resolution is a precondition for accurate time synchronization.

Stability of the System Time

Usually the system clock is implemented as a counter chain driven by an oscillator, and a certain number of counts are interpreted as 1 s. If the system clock drifts, then the reason may be that the oscillator frequency has a mean offset from its nominal frequency and in addition varies due to:

- Variations in the ambient temperature caused by:
- Daily temperature changes
- Air condition kicking in and out
- CPU temperature varying with system load

- Gravitational or acceleration effects in moving systems as well as relativistic effects, e.g., in satellites

The more the oscillator frequency is off, the faster the time drifts and the larger is the compensation to be applied by a time synchronization software.

Possible Ways to Adjust the System Time

Time synchronization software can adjust the oscillator frequency in a way that both the clock drift and a time offset are minimized if the hardware supports this, e.g., if the oscillator which drives the counter chain for timekeeping has a control voltage input.

However, on many systems like standard PCs, it isn't possible to tweak the real frequency of the oscillator assembled on the mainboard, but another possible way is to change the number of clock counts per seconds, e.g., by adjusting a counter compare or reload register. Of course, corrections can only be made in integral steps of the clock cycles, so a higher clock frequency allows a finer adjustment.

Characteristics of the Time Synchronization Software

The time accuracy which can be achieved depends strongly on the implementation of the time synchronization software running on the client.

- Which time resolution is supported in computations? If the difference between the reference time and the local system time can only be determined coarsely, e.g., in units of whole seconds, then the adjustment can't be very accurate.
- Are transmission delays compensated as accurately as possible? If not, then there may be systematic time errors which aren't visible at the first glance.
- How is the client time adjusted? Set periodically? Smoothly? If corrections are only applied by setting the time in periodic intervals, then the time drifts more or less during an interval, depending on the oscillator's mean frequency offset.
- Is the system clock drift compensated? Even if time adjustments are applied smoothly, without stepping the time, it is important that the clock drift is also determined and compensated as good as possible so that time offsets don't even evolve.

Handling of Leap Seconds

Leap seconds are scheduled by the *International Earth Rotation and Reference Systems Service* (IERS) at certain points in time to keep the Coordinated Universal

Time (UTC) in accordance with the Earth rotation (IERS 2016a). Typical intervals between leap second events have ranged from 12 or 18 months up to even 7 years, which was the longest interval between 2 leap seconds (IERS 2016b).

Some reference time sources like GPS receivers can provide a leap second announcement, and network protocols like NTP and PTP can pass the announcement on to their clients. It depends on the software running at the client side how a leap second is applied. On Unix-like systems, the leap second is simply passed down to the kernel, which steps the time back by 1 s to insert a leap second. This may affect, e.g., database applications which expect the system time always to increase continuously.

Windows itself doesn't handle leap seconds at all, but the *ntpd* port for Windows contains a work-around which slews the time softly across the leap second, without any steps (Meinberg 2016).

Virtualized Systems

Modern systems are often running in virtual machines (VMs) rather than directly on physical servers. This has advantages for the management of these machines, but it also means that even timer ticks for each individual VM are emulated by the host system instead of being derived directly from a physical oscillator and counter chain.

So the basic stability of the system time inside a VM depends on the implementation of the virtualization software. If the virtualization software doesn't provide the guest systems running in the VMs with a stable timing, then time synchronization software may not even be able to correct this. Also, if a VM is suspended or moved to a different physical machine, the system time may be temporarily stopped, so that a time step may occur after the operation has finished (VMware 2011).

Time Distribution

There are different ways to distribute the time from a reference time source locally. Figure 38.6 shows some examples.

A single GNSS/GPS antenna can either provide a single receiver with the satellite signals, or an antenna diplexer can be used to distribute the satellite signals to several receivers. The advantage of using a diplexer is that only a single antenna cable needs to be wired to the roof, which can be very expensive depending on the owner of the building. The disadvantage is that all GNSS receivers depend on a single antenna, so there's no redundancy.

There are stand-alone GPS receivers, GPS PCI cards which can be plugged into servers to make a very accurate time available to the server and NTP/PTP time

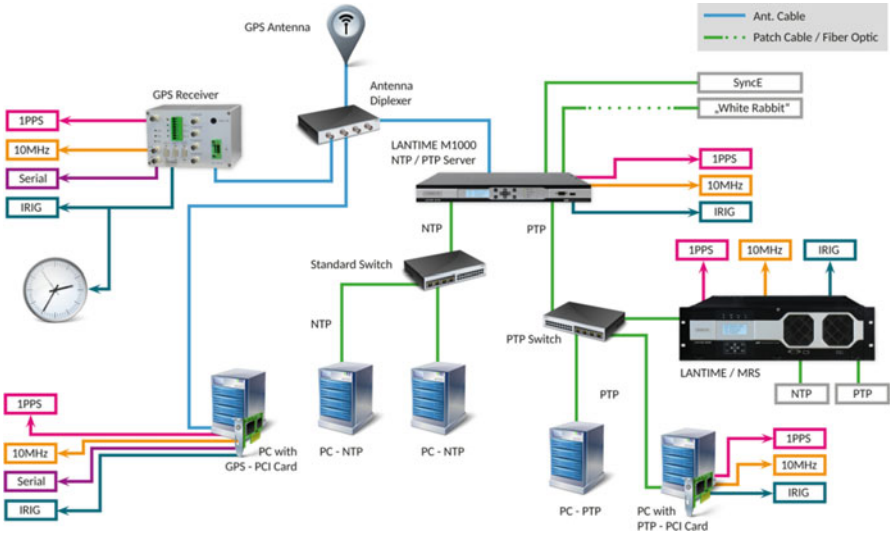


Fig. 38.6 Example of time signal deployment

server appliances which get the time from a built-in GPS receiver and make the disciplined time available via some network protocols, e.g.:

- NTP, which provides good accuracy for clients with moderate accuracy requirements
- PTP, which provides highest accuracy in a LAN for special requirements
- "White Rabbit," which is based on PTP and provides extreme accuracy but requires very special hardware

Local distribution of very accurate time via PTP has the advantage that standard patch cables can be used instead of special antenna cables, but on the other hand, PTP-aware switches are required for the PTP protocol to yield highest accuracy. PTP PCI cards can get a very accurate time into a server, similar to GPS PCI cards.

Most of the specialized timing cards and devices can output a number of different hardware signals, e.g.:

- 1 PPS signal
- 10 MHz frequency
- Serial time strings
- IRIG and similar time code signals

which can in turn be used to discipline the time of other device. For example, an IRIG time code signal can be used to synchronize special measurement devices or simple wall clocks driven by a time code.

Network Time Synchronization Protocols

The *Network Time Protocol* (NTP) (Mills et al. 2010) and *Precision Time Protocol* (PTP, IEEE 1588) (The Institute of Electrical and Electronics Engineers, Inc. 2002) are network protocols which can be used to synchronize the time across a local area network (LAN) or wide area network (WAN).

NTP was the first time synchronization protocol which supported high-resolution time stamps when it was introduced in the 1980s by Dave Mills. The current protocol level is NTPv4, which is compatible with earlier versions. NTP can typically achieve a few milliseconds of accuracy over Internet/WAN connections, and even 100 µs or better on a LAN, but this *can't be guaranteed* and depends on some characteristics of the overall system. There are a couple of major implementations available:

- *ntpd*, the NTP reference implementation introduced by Dave Mills and currently maintained by the NTP project, is a free software which supports a huge number of radio clocks, GPS receivers, ACTS service, etc. to be used as reference time source in server mode and can yield an impressive accuracy in client mode, without requiring special network hardware. This software package is widely used on Unix-like systems and has also been ported to Windows (The NTP Public Services Project 2016).
- *chrony* is a newer implementation of an NTP daemon which runs on Linux, FreeBSD, NetBSD, Mac OS X, and Solaris. One of the main goals of this open-source software is fast settling of the control loop for the system time (chrony homepage 2016).
- *OpenNTPD* is a free NTP implementation maintained by the OpenBSD project (OpenNTPD homepage 2016).
- *w32time* is the proprietary time service from Microsoft, which is shipped with Windows and also uses NTP as the network protocol. Usually it yields less accurate time than the *ntpd* port for Windows (Microsoft Developer Network 2007).

PTP is a newer network protocol which can typically yield a few nanoseconds of accuracy on local networks, but *only* if specific network hardware is being used, i.e., *all* involved network nodes support hardware time stamping, and specifically all routers and switches provide extra support for PTP. If no special hardware is used, then PTP has the same limitations as NTP. Version 1 of the PTP protocol was published as IEEE 1588 in 2002, and in 2008 the newer version 2 was released, which unfortunately isn't compatible with version 1.

Popular implementations are:

- *PTPd*, an open-source software package maintained at GitHub, which can be used under Linux and FreeBSD (PTPd official source 2016)
- *LinuxPTP*, a free implementation maintained by the Linux PTP project, which has been specifically designed to be used with a Linux kernel (The Linux PTP Project 2016)

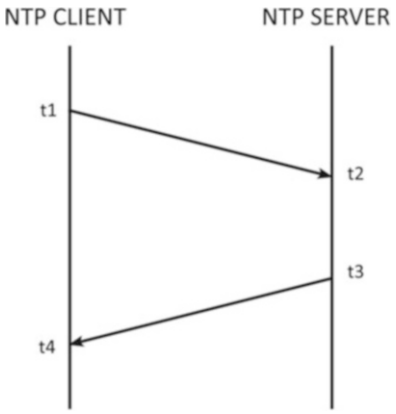


Fig. 38.7 NTP packet exchange

A new approach is the *White Rabbit* project which runs the PTP protocol over *Synchronous Ethernet* (SyncE) on dedicated fiber optic connections, with special hardware, and thus yields sub-nanosecond accuracy (Open Hardware Repository 2016).

Both NTP and PTP are simple at the network packet level, and the achievable accuracy depends mainly on the evaluation of the network packet by the *client software*.

General Network Time Transfer

If time servers simply send time to their clients, e.g., as broadcast packets, then clients which receive a packet don't know how long the packet has been traveling before it was received. Thus, most protocols try to determine and compensate the packet propagation delay on the network. Figure 38.7, for example, shows the packet exchange for an NTP client querying the time from a server.

- t1: Client sends request packet to server.
- t2: Server receives request packet from client.
- t3: Server sends reply packet to client.
- t4: Client receives reply packet from server.

So there are four time stamps (t1–t4) from one packet exchange which need to be evaluated by the *client*. Care must be taken that the time stamps taken on the server (t2, t3) are from the server's system time, while the time stamps taken at the client side (t1, t4) are taken from the client's system time, which may have an unknown offset to the server's time. So the client has to do some math on the four time stamps to separate the time difference from the packet propagation delay.

This works well if the network delays in both directions are the same, but if one of the packets takes longer to be delivered than the other one, then this causes an error in the calculation of the time offset, so the client should do some filtering to

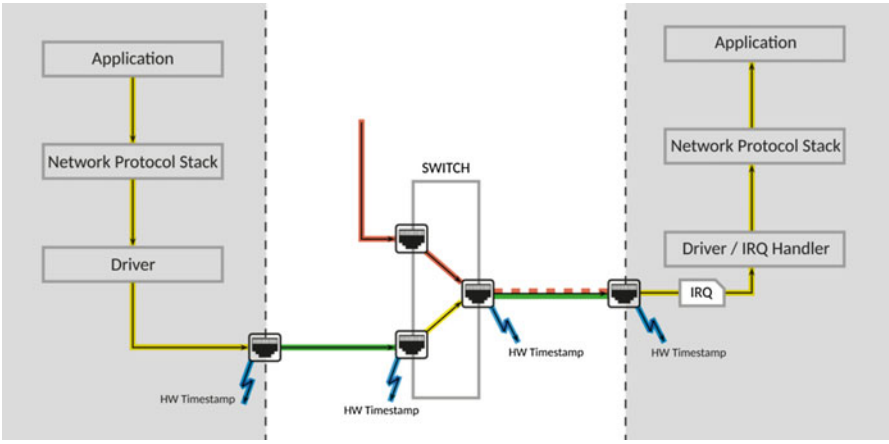


Fig. 38.8 Network packet latencies

determine the mean packet propagation delay and discard packets which have been traveling too long.

If a part of the network route is an asymmetric line with different bandwidths in both directions (e.g., ADSL), then *all* packets are affected in the same way. This causes a systematic time error which can't be determined by a client and isn't reflected in the time offset reported by the client.

Network Packet Latencies

When a packet is sent from one node to another, then there are different delays which need to be determined and compensated, if possible. Figure 38.8 shows a schematic.

Once an application has sent a packet, the packet goes down through the network stack to the driver for the network interface card (NIC), which sends the packet out on the wire. How long this takes depends on the implementation of the network protocol software, the CPU power, and the system load.

The packet then travels across a network cable, eventually to a switch, where it comes in at one network port and goes out at another port. If the outgoing port is then busy sending a different packet or if the incoming and outgoing ports have different link speeds, the packet is delayed inside the switch for an unknown time interval.

After the switch, the packet travels again across a cable to another switch or router, which may cause additional delays, or finally to the destination computer.

When the packet comes in from the wire at the target computer, the NIC generates an interrupt request (IRQ), and after some interrupt latency, the NIC driver's interrupt handler is activated to pick up the packet.

Modern NIC chips have a feature called *interrupt coalescence*. This means the NIC chip has a FIFO buffer which can hold several packets, and the chip only generates an IRQ after several packets have been received. This feature is very nice to achieve high network throughput, but it is very bad for timing applications since the interval until a packet is picked up after it has come in from the wire can widely vary.

Finally the packet is sent up the protocol stack to the application which waits for it. Again, this execution time depends on the implementation of the network protocol software, the CPU power, and the system load of the target machine.

So the only constant delays on the whole way are the delays on the cable. The approach PTP is using to eliminate the unknown delays is to support hardware time stamping on every node (marked by flashes in Fig. 38.8) and send the captured time stamps to the client so the client can do the computation required to eliminate the unknown delays. Obviously this also needs to be implemented on routers and switches, since otherwise there were still unknown delays in the network path.

In case of a wireless (Wi-Fi) connection things are even worse since the transmission delay over the air varies strongly with the signal level, bandwidth, and usage of the Wi-Fi connection.

Network Time Client Requirements

As shown in section "Network Packet Latencies", the network delays are not constant, so some filtering is required on the client, or hardware time stamping has to be used to eliminate the network jitter.

So it is important to keep in mind that the accuracy which can be achieved on a client doesn't only depend on the accuracy of the server but also strongly on the client hardware and software implementation and of course on the network characteristics.

With most protocols, only the *client* can determine and compensate the transmission delays. There are also exceptions where the *server* does the delay compensation, but this requires that the server keeps track of the client's state and applies an individual correction for each client. For example, the ACTS telephone service works this way, but this is not feasible, e.g., for NTP or PTP servers with a huge number of clients.

Generally, when talking about NTP or PTP, we need to distinguish between the *name of the protocol* and its *implementation*.

PCI Access Latencies

If a PCI card is installed in a computer and the CPU is going to read a time stamp from the card, it must be kept in mind that the PCI bus on the mainboard is basically similar to a small network, so there are similar latencies. Figure 38.9 shows the basic structure.

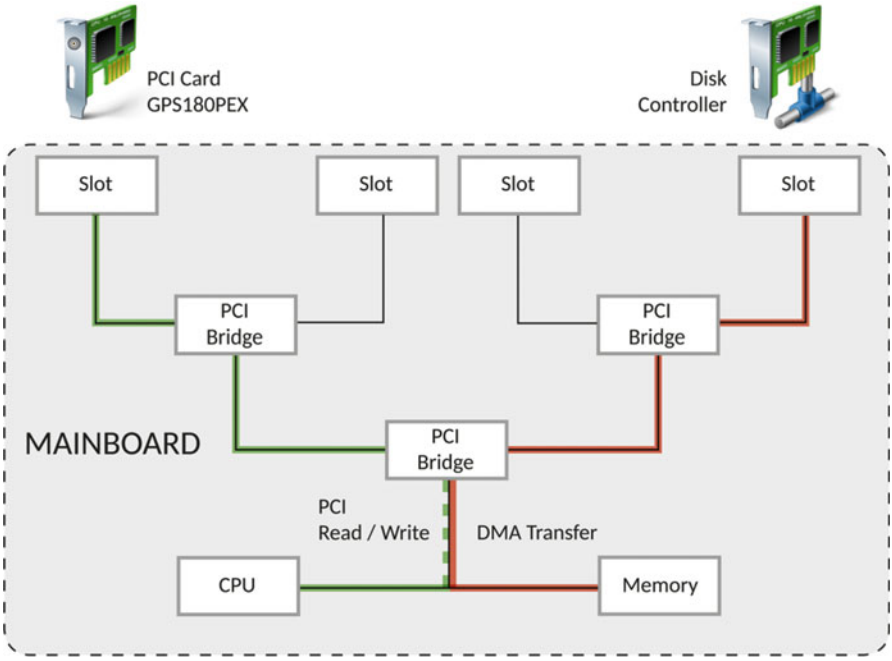


Fig. 38.9 PCI bus structure

Each PCI bus slot, and each PCI device assembled on the mainboard, is connected to the CPU via one or more PCI bridges. For the consideration here, the function of these PCI bridges is comparable to network switches.

When the CPU is going to read some data from the PCI card in the left slot, then this may succeed immediately and quickly, except in case the bus is actually busy. For example, if there is an ongoing DMA transfer from the disk controller to the system memory, then the shared part of the bus between the CPU and the first bridge is busy, and the read access from the CPU is delayed by the PCI bus arbitration until the bus is free again. This is unnoticeable by the application which simply lets the CPU execute a *read* command.

So the execution time required to read a time stamp from a PCI card also depends on the bus architecture on the mainboard (i.e., the number and type of the bridges), and it may vary depending on the bus arbitration or the CPU being interrupted when it is going to execute the *read* command.

So a single access to read a 64-bit time stamp from a PCI card can take a few microseconds to execute. This doesn't sound much, but it significantly limits the accuracy available to an application if the time of the PCI card is accurate to a few nanoseconds.

Hardware Signals

Devices like radio clocks and GNSS receivers often provide hardware signal outputs which can be used to synchronize peripheral devices, or distribute the time or frequency locally to other devices. There are different types of signal with different characteristics.

The accuracy of an output signal usually depends on the time quality and accuracy provided by the specific time source. For example, output signals generated by long wave receivers are usually less accurate than output signals from GNSS receivers.

Generally it has to be distinguished between digital signals with rectangular shape and analog signals which are often sine wave signals with a specific frequency. In some cases, a sine wave signal is also used as a carrier for a digital signal or time code.

In industrial environments, or if signals have to be forwarded over long distances, it is useful to take care that the transmitter and receiver are electrically insulated, so that the devices can't be damaged by ground loops or similar. For most signal types, specific converter circuits are required for insulation, for example, transducers for analog signals or optocouplers or fiber optic converters for digital signals. These converters can cause signal delays which may need to be determined and compensated if high accuracy is required.

Digital signals usually have a rectangular shape with specified voltage level swings, e.g., 0 V through 3 V, or 0 V through 5 V. In regard of the signal level, these are asymmetric signals, so they are susceptible to electrical noise and thus should only be used for short distances, e.g., a few meters.

The noise immunity can be increased if RS-232 level converters are being used at the ends of a transmission line, and thus a voltage range of ± 3 to ± 12 V can be used for the transmission. However, many industrial RS-232 converters produce output slopes which are rising only slowly and thus introduce a significant signal delay and jitter at the receiving side.

A better approach is to use RS-422 or RS-485 converters at each end of a transmission line. These are symmetric signals and thus have a much better noise suppression than RS-232, so they can be used for much longer distances than the signals discussed before. These signal types also allow higher data rates than, for example, RS-232, and thus insert less jitter. The main difference between RS-422 and RS-485 is that RS-422 is specified for point-to-point connections, i.e., one transmitter feeds one receiver, while RS-485 can be used for bus topologies where one transmitter can talk to several receivers.

All these signal types may still require electrical insulation, which may be achieved by optocouplers inserted at least at one end of the transmission line or alternatively by using fiber optic transmission.

Fiber optic transmission is not susceptible to electric noise at all and has a "built-in" insulation, so it can be used over very long distances, e.g., up to several kilometers, depending on the type of fiber optic. It also supports very high data

Fig. 38.10 Unmodulated and modulated time code signals

rates, which means it doesn't cause much jitter to the transmitted signals. However, eventually a signal delay by the transceivers has to be taken into account.

A 10 MHz frequency is often derived directly from a built-in high-quality oscillator, which may even be disciplined by an input signal. The output signal may be square wave, in which case the same transport media can be used as described above, or sine wave, in which case analog transmission has to be chosen, including transducers if insulation is required.

A pure frequency can neither transport the absolute date and time nor can it mark any second boundaries like a PPS signal, but it can be used to provide other devices with a stable reference frequency and thus avoid clock drifts.

IRIG/IEEE time codes are a combination of digital pulses and absolute time information. There are specifications for unmodulated (DC level shift, DCLS) and modulated signals. Modulated signals consist of a sine wave carrier with a particular frequency onto which the DCLS signal is modulated. They are easy to handle, can be distributed from one output to several inputs, and even be recorded, e.g., on magnetic tape. Since it's an AC signal, there's no need to care about polarity.

Unmodulated (DCLS) codes can be distributed over the same physical ways as the other digital signals mentioned above, so care must be taken that the input and output drivers at both sides are laid out for the same signal type, polarity, and signal level.

Figure 38.10 shows an example for an unmodulated and modulated time code signal. You can easily see that it's easier to decode a certain point in time from a

particular DCLS slope than from the associated modulated signal. That's why DCLS codes normally yield more accuracy than modulated codes. The difference may only be a few microseconds, though.

There are different IRIG code frame formats which have been defined to transport different amounts of information. All code formats provide at least the day of year and the time of day. However, many commonly used IRIG code formats don't provide a year number, a UTC offset, or daylight saving and leap second flags. Without a year number, the conversion from the transmitted day of year to a calendar date is ambiguous. The 29th of February only exists if the year is a leap year. Without a UTC offset, the receiver can't distinguish if the transmitted time is the local time of some time zone, or UTC, and without a leap second announcement flag, the receiver is unable to handle a leap second at the exact point in time.

The original IEEE 1344 standard and its revised version which has been renamed to IEEE C37.118 are compatible with the popular IRIG-B002/B122 code but provide all the additional information in a bit field of the data stream which is reserved in B002/B122.

So usage of one of the IEEE codes is highly encouraged because these codes allow unambiguous decoding of the time code by receivers which know how to interpret the extended information, while legacy B002/B122 receivers ignore the extended information and can still decode the day of year and time of day.

Care must be taken with the IEEE codes that the way the UTC offset has to be applied has been reversed from IEEE1344 to C37.118, so some time code generators and receivers explicitly distinguish between both code formats.

Example Measurement Results

Some example measurements with *ntpd* using different time sources show how the resulting accuracy depends on the selected time sources.

Figure 38.11 shows the loopstats recorded by ntpd on a laptop running Linux with a Meinberg TCR51USB IRIG time code receiver as reference time source.

Even though the accuracy of the reference time provided by the IRIG receiver is in the range of a few microseconds, the time offset determined by ntpd is in a range about ± 20 µs. These results are still pretty good under the given conditions since the jitter caused by the USB bus is reduced by ntpd's filtering.

Figure 38.12 shows a loopstats file from a Linux PC with a built-in Meinberg GPS180PEX card and the same scale used for the time offset as in Fig. 38.11.

The results are much better than those shown in Fig. 38.11, but no details can be seen at this time offset scale, so Fig. 38.13 shows the same data with higher resolution.

This plot is from the same data as Fig. 38.12, but it shows that the determined time offset is in the range of a few microseconds only during the measurement interval.

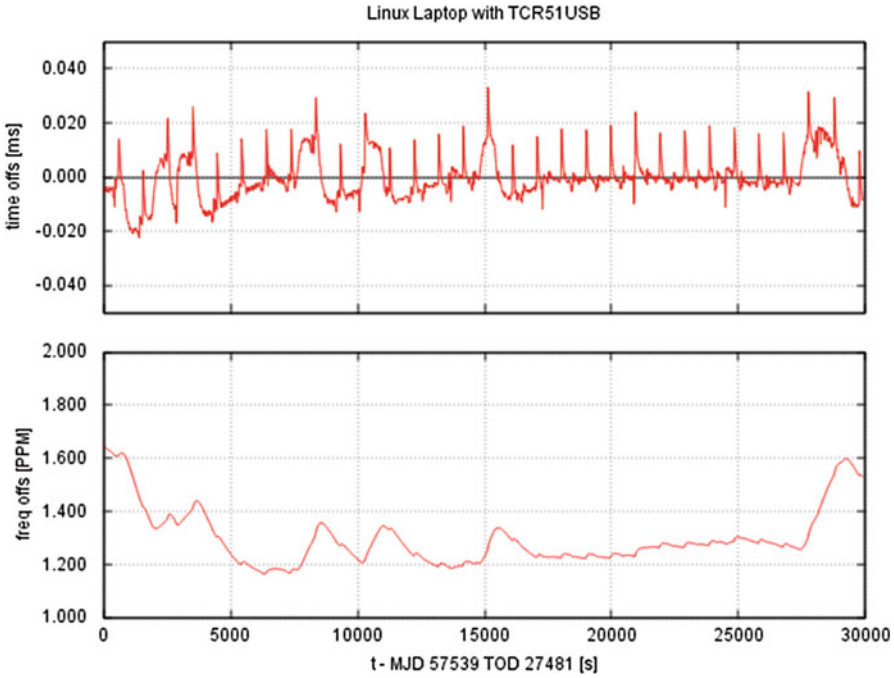


Fig. 38.11 Linux laptop with TCR51USB IRIG receiver

There are no significant variations in the frequency offset graph, and the absolute value of the frequency offset reflects the native frequency offset of the oscillator on this particular mainboard, which has been compensated by ntpd.

However, if we take a look at a longer interval, then we see that the time adjustment results aren't always as accurate as shown in Fig. 38.13. Figure 38.14 shows plots of the full data set which has been recorded over approximately 4 days.

The frequency offset plot now reflects cyclic frequency variations which are due to the daily temperature changes in the room where the computer is located.

Whenever the temperature changes, the frequency of the oscillator on the mainboard also changes, which lets the system time start to drift. When ntpd notices the increasing time offset, the control loop is adjusted accordingly to minimize the time offset again, which results in a variation of the frequency offset reported in the loopstats files.

If no directly attached reference time sources are available and the time is synchronized only across the network, then the network latencies and jitter affect the achievable accuracy. Figure 38.15 shows the plot of another loopstats file recorded on a Linux laptop, where servers of the public NTP pool are used as time sources. The laptop is connected to the Internet via Wi-Fi, which makes the expected results even worse.

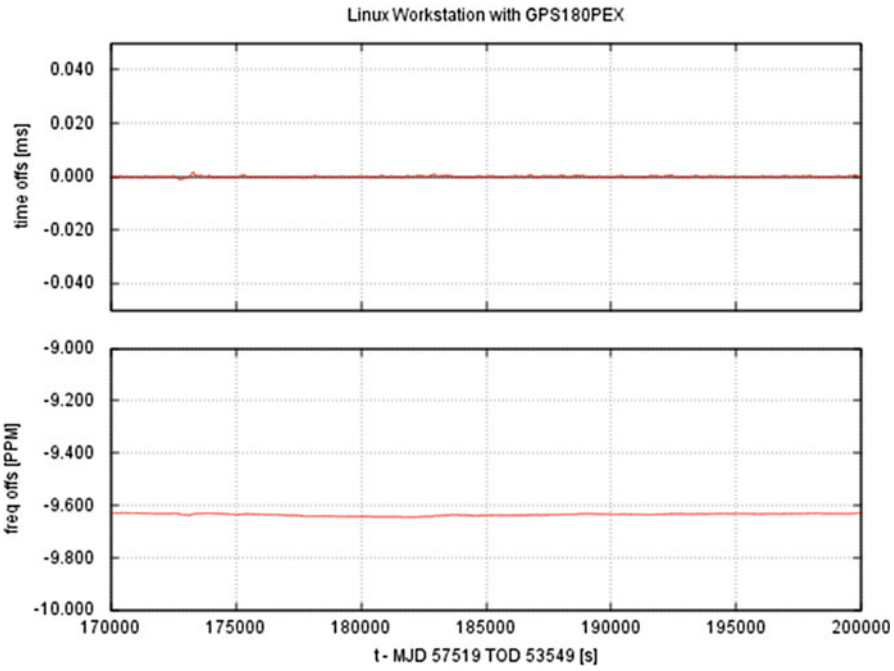


Fig. 38.12 Linux PC with GPS180PEX GPS PCI card

The time offset scale in this plot is now ± 10 ms, not microseconds. Anyway, the achieved accuracy is still good under the given conditions.

Comparing Offsets Measured by ntpd

The NTP daemon (ntpd) can easily be used to measure the offset of remote NTP nodes with moderate accuracy.

Ideally there should be a local reference time source which provides accurate time. In the example here, we use the Linux laptop again with an IRIG receiver as reference time source, and let ntpd measure the time offset of another NTP server on the local LAN. A few lines in the NTP configuration file (ntp.conf) are sufficient for this configuration:

```
server 127.127.28.0 iburst minpoll 4 maxpoll 4
fudge 127.127.28.0 refid shm0
server 192.168.0.2 iburst minpoll 4 maxpoll 4 noselect
```

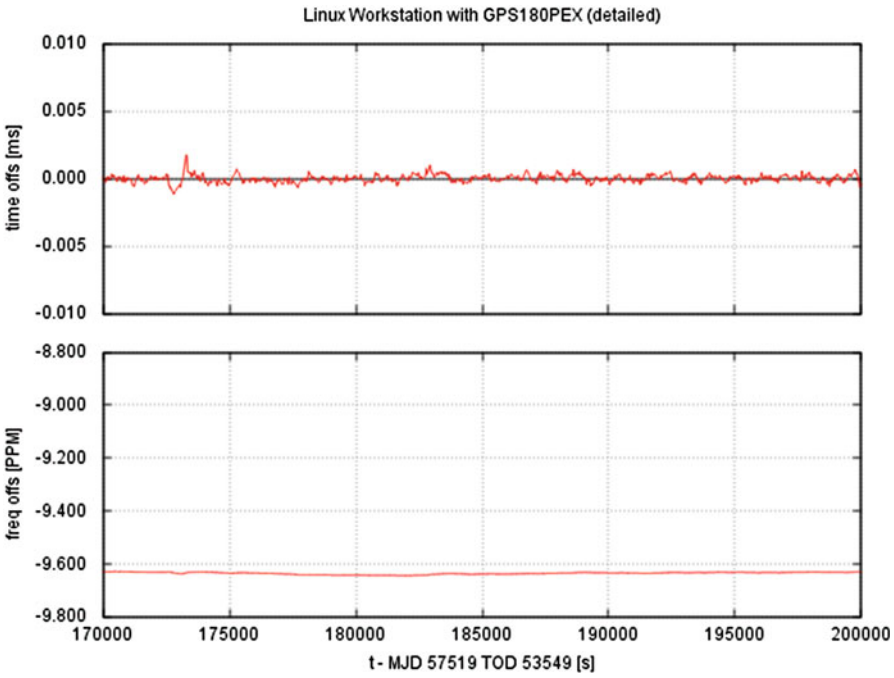


Fig. 38.13 Linux PC with GPS180PEX GPS PCI card, detailed

The first two lines referring to 127.127.28.0 tell ntpd that the reference time is supplied via its shared memory driver, which should be displayed as *shm0*.

Line 3 specifies the NTP server to be tested. The important keyword here is *noselect*, which tells ntpd to query the time from the server and evaluate the results (i.e., poll the server), but don't use the results from this server to discipline its own system time.

The measurement results can, for example, be checked by periodically running the command *ntpq –p* which shows an output like:

```
     remote           refid      st t when poll reach   delay   offset  jitter
==============================================================================
*SHM(0)          .shm0.          0 l    6   16   377    0.000    0.001   0.002
 192.168.0.2     .MRS.           1 u    3   16   377    0.187    0.013   0.065
```

The delay, offset, and jitter figures shown in the output are in milliseconds and are related to the own system time. The first line of the printed table refers to the configured reference time source, SHM(0) in this case, which has only 1 µs offset.

The second line which starts with the IP address of the NTP server we want to test reports the measurement results for that server. The computed offset is 13 µs in

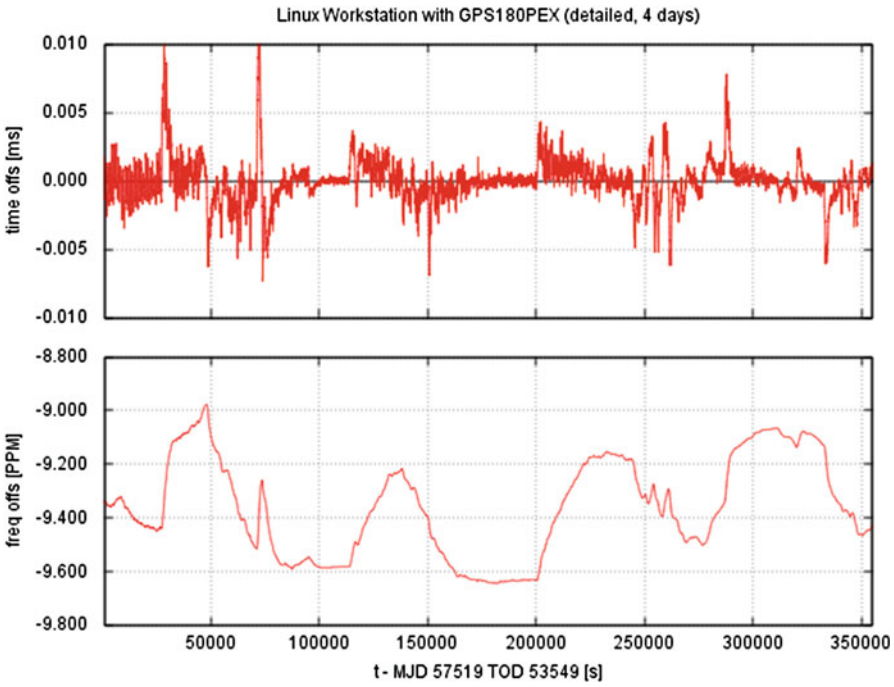


Fig. 38.14 Linux PC with GPS180PEX GPS PCI card, detailed, 4 days

this example, but the jitter observed from several polls is 65 µs. This is due to the network jitter and can be used as an indicator for the range of error to be taken into account for the computed time offset.

Conclusion

As a conclusion, we can say that the time synchronization accuracy depends on many facts like the used protocol or signal type, the latencies and jitter between time source and client, and which effort is made at the *client* side to minimize potential errors.

Also the hardware characteristics (mainboard, CPU) and software characteristics (operating system, time synchronization software, virtualization) at the client side can strongly affect the achievable accuracy.

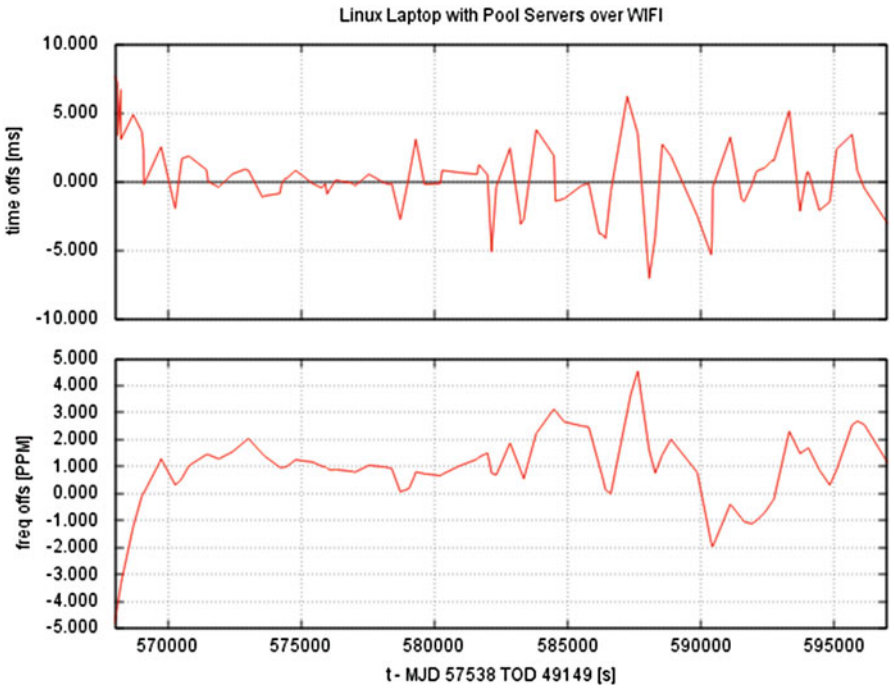


Fig. 38.15 Linux laptop, pool servers over Wi-Fi

Different applications have different accuracy requirements, and higher accuracy requirements need higher effort to achieve the goals, which may result in more expensive solutions. So it is very important to select the most appropriate solution for a specific application.

References

chrony homepage (2016). https://chrony.tuxfamily.org. Accessed 30 July 2016

IERS, Earth Orientation Center (2016a). https://hpiers.obspm.fr/eop-pc/index.php. Accessed 30 July 2016

IERS, Leap Second (2016b). http://hpiers.obspm.fr/eop-pc/earthor/utc/leapsecond.html . Accessed 30 July 2016

Meinberg, Leap Seconds and how they are handled by Meinberg Devices and NTP (2016). https://www.meinbergglobal.com/english/info/leap-second.htm. Accessed 30 July 2016

Microsoft Developer Network, What is Windows Time Service? (2007). https://blogs.msdn.microsoft.com/w32time/2007/07/07/what-is-windows-time-service. Accessed 30 July 2016

Mills et al., IETF RFC 5905, Network Time Protocol Version 4: Protocol and Algorithms Specification (2010). https://www.ietf.org/rfc/rfc5905.txt. Accessed 30 July 2016

NIST, WWVB Coverage Area (2016). http://tf.nist.gov/stations/wwvbcoverage.htm. Accessed 30 July 2016

NTP Documentation, NIST/USNO/PTB Modem Time Services (2012). http://doc.ntp.org/4.2.8p8/drivers/driver18.html. Accessed 30 July 2016
Open Hardware Repository, White Rabbit (2016). http://www.ohwr.org/projects/white-rabbit. Accessed 30 July 2016
OpenNTPD homepage (2016). http://www.openntpd.org. Accessed 30 July 2016
PTPd official source (2016). https://github.com/ptpd/ptpd. Accessed 30 July 2016
The Institute of Electrical and Electronics Engineers, Inc., *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*. IEEE Std. 1588–2002 (New York, 2002). ISBN 0-7381-3369-8
The Linux PTP Project (2016). http://linuxptp.sourceforge.net. Accessed 30 July 2016
The NTP Pool Project (2016). http://www.pool.ntp.org/en/. Accessed 30 July 2016
The NTP Public Services Project (2016). http://support.ntp.org. Accessed 30 July 2016
VMware, Timekeeping in Virtual Machines (2011). http://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/techpaper/Timekeeping-In-VirtualMachines.pdf. Accessed 30 July 2016

Chapter 39
Time Critical: Contesting the Measure of the Now

Daniel Wiley

Abstract It is increasingly well understood that transformations in the concept of time, as well as techniques of time measurement and distribution, around the turn to the twentieth century contributed to the formation of entrenched divisions between the Sciences and the Humanities. In general, where the Sciences have conceived of time as a function of conventionally defined quantitative measure, the Humanities have consistently regarded human experience as an essential element in any adequate view of time.

Using Media Studies as an example, this paper reflects on the status of these divisions amidst the postwar rise in interdisciplinarity. From its beginnings, Media Studies has been concerned with how scientific and technological objects and processes impact individual and collective temporal experiences and behaviors, as well as formal and everyday epistemologies of time. Despite these broadened horizons, however, Media Studies remains largely mired in the Science-Humanities divide. I argue that this is due, in large part, to an overreliance on philosophical concepts and positions that are foundational to this divide and a consequent reluctance not to subordinate the scientific view, from the outset, as derived from the "real" human experience of time. I propose that this imposition of hierarchy is misguided and that measurement is fundamental to the human experience of time. In conclusion, I draw on recent resources from the history and philosophy of Science and Media Studies to suggest how a more fruitful cross-pollination of knowledge from the Sciences and Humanities might be approached and applied to our understanding of time-based media.

Keywords Time • Measurement • Science • Humanities • Twentieth-century history • Philosophy • Media studies

It is increasingly well understood that transformations in the concept of time, as well as techniques of time measurement and distribution, around the turn to the

D. Wiley (✉)
Department of Media, Culture, and Communication, New York University, New York, NY 10003, USA
e-mail: dmw382@nyu.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_39

twentieth century contributed to the formation of entrenched divisions between the Sciences and the Humanities. In general, where the Sciences have conceived of time as a function of conventionally defined quantitative measure, the Humanities have consistently regarded human experience as an essential element in any adequate view of time.

Chapter 40
Timescale Pluralism and Sciences of Time

Kevin Birth

Abstract The history of the science of time often focuses on the techniques and technologies that led to our present practices but neglects other approaches to timekeeping that were discarded, neglected, or peripheral to how we currently reckon time. Yet, there have been contexts in which multiple timescales coexisted, even if such instances of timescale pluralism have now been mostly forgotten. This presentation explores timescale pluralism and timescale uniformity in late medieval York, early modern Nuremberg, mid-twentieth century studies of the biting behavior of mosquitoes in Trinidad, and current primatology. These different cases reveal the benefits and pitfalls of timescale pluralism versus timescale uniformity and suggest a rethinking of the history of the science of time away from narratives that emphasize function and progress to those that explore politics, contested choices, and the conceptual costs of particular decisions.

Keywords Timescale pluralism • History of timekeeping • York • Nuremberg • Regiomontanus • Mosquitos • Baboons

K. Birth (✉)
Department of Anthropology, Queens College, CUNY, Flushing, NY 11367, USA
e-mail: kevin.birth@qc.cuny.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_40

Chapter 41
Liberating Clocks: Exploring Other Possible Futures

Michelle Bastian

Abstract This paper suggests that alongside concerns regarding the future of precision timekeeping, there is also a need to discuss the future of timekeeping in social life. Pointing to the way that maps are created to fit a range of different user needs, I ask whether methods of telling the time might also be made more open to experimentation and redesign. In order to provide examples of how this could be done, I draw together a range of projects by artists and designers who are using clocks in unexpected ways. Looking at examples that particularly focus on social aspects of environmental issues, I show how the clock can be a useful tool for highlighting alternative ways of keeping time. Drawing on work in temporal design, developed in collaboration with designers Larissa Pschetz and Chris Speed, this paper suggests a new development in the field of horology, towards a critical horology that emphasises the political, social and environmental aspects of timekeeping.

Keywords Horology • Clocks • Design • Environment • Critical horology

Introduction

The 2016 Science of Time conference, subtitled 'Time in Astronomy and Society, Past, Present and Future', has offered a range of insights into the ways that timekeeping might respond to changing user needs. These changes were exemplified in the call for papers by references to shift in standards, in methods of distribution and in the definitions of units and timescales. As a philosopher interested in the role of time in social life, my paper offers another possibility for the future of timekeeping, namely, that clocks and clock time might respond, not only to the changing needs of science but also the changing needs of society more broadly. I first suggest that the humanities have much to learn from discussions

This paper draws on ideas from a longer paper on 'Liberating Clocks' which is forthcoming in the journal *New Formations*.

M. Bastian (✉)
Edinburgh College of Art, University of Edinburgh, 20-22 Chambers Street, Edinburgh EH1 1JZ, United Kingdom
e-mail: michelle.bastian@ed.ac.uk

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_41

occurring within the community gathered together by 'the science of time', particularly my own field of continental philosophy. However, I also suggest that deliberations on the science of time might be usefully broadened out through attention to work arising from art and design. In particular, I showcase examples of 'clocks' that highlight alternative ways of measuring and telling the time and thus open up the field of timekeeping to broader applications and experimentations.

A Single Objective Clock Time?

Within my own discipline, specifically continental philosophy, clock time is generally thought of as being external to society and thus not particularly capable of addressing social issues of core interest. A good example of this can be seen in David Couzens Hoy's *Time of Our Lives*, where he writes that the 'term "time" can be used to refer to universal time, clock time, or objective time. In contrast, "temporality" is time insofar as it manifests itself in human existence' (Hoy 2009, p. xiii). To unpack this a little, in listing synonyms for 'time', clock time is assumed to be roughly equivalent to 'universal time' or 'objective time'.[1] Further, clock time is placed in opposition to the kind of time that appears within 'human existence' or what Hoy refers to as 'temporality'. Importantly, given that continental philosophers are particularly interested in temporality, rather than time, this means that clocks and clock time are, more often than not, placed outside of their sphere of main concern.

However, this kind of philosophical approach has much to learn from time metrology, time standardisation and horology. While the norm within continental philosophy is to assume that clock time refers to a single, measurable, objective 'time' and is thus incapable of dealing with more complex approaches to time, a brief review of work presented at this conference significantly challenges this. Daniele Rovera described UTC as something that is 'produced', rather than simply measured, thus suggesting a larger role for social and technical processes in timekeeping than continental philosophers have acknowledged. Further, Dennis McCarthy highlighted the multiple timescales available for scientific use and described them as a 'family', even providing a family tree. Speaking even more directly to the incorrect assumption that there is a single way of measuring time, Danny Hillis discussed the multiple kinds of 'time' (e.g. derived solar time, monotonic corrected time, planet display time) that will feed into the time indicated by the Clock of the Long Now. His comment that this complexity was rarely acknowledged by the general public was met by audible and knowing agreement from attendees and chimes with my experiences in my own discipline. Even what was called a clock was much more open and nuanced, since speakers talked of stars,

[1]Note that Hoy does not use 'universal time' in any technical sense here but rather refers to a somewhat vague idea of astronomical time.

radioactive decay, pulsars and black holes all acting as 'clocks'. Each one of these claims challenges continental philosophical 'common sense' and could potentially encourage a deeper curiosity within the discipline about how time telling should be understood.

Work presented at the conference also offered further challenges to the assumption that the tools and methods for telling time are separate from human existence, as Hoy supposes. For example, Sara Schechner demonstrated the role of power and authority in the creation of sundials, arguing that they can be read for so much more than the time they tell (e.g. Schechner 2001). McCarthy also recalled previous conversations held within the time metrology and horological community about the ways that time is shaped by political authority. This included stories of Ancient Romans bribing the Pontifex Maximus to make a year shorter or longer (Hewitt Key 1875, p. 230), while in Egypt, Pharaohs were required to swear to not make any changes to the calendar when they took the throne (Richards 2000, p. 110). A further example came from Steve Allen's presentation on time zones, where the production of the tz database brought to the fore the difficulties of responding to technical and political challenges within the same system. Thus, for those interested in the future of timekeeping in social life, the approaches taken to time within time metrology and horology are worthy of far greater attention.

Prioritising Precision Timekeeping?

Even so, the rest of my paper suggests that just as continental philosophers can learn much from scientific approaches, these scientific approaches might also be usefully expanded by looking at the interventions of artists and designers into questions of time measurement and timekeeping. More specifically, I suggest that these interventions question the priority given to the development of greater precision and accuracy. Throughout the history of timekeeping, a history that was reflected in the conference, it is generally accepted that progress towards these aims has been achieved by improving methods for dealing with external perturbations. Indeed, in a number of presentations, speakers talked of needing to 'immunise', 'insulate' and 'isolate' clocks from the effects of social and material processes. Of particular resonance for a humanities scholar was William Andrewes' discussion of the design of the Harrison clocks and the observation that a heavy pendulum is problematic, because it 'remembers everything' (the example given by Andrewes was the vibrations from a horse and carriage going past outside). A light pendulum, on the other hand, 'forgets' more easily and is less affected by its surroundings. Further, speakers argued that other potential problems to do with cosmological time or time signalling led to a need to make assumptions about symmetry and homogeneity, either for the sake of 'progress' in cosmological time telling (Impey) or for more pragmatic reasons of having a time service that works within acceptable error margins (Levine).

My proposal, however, is that we see the emphasis on concepts such as accuracy, reliability, homogeneity and precision as one particular focus and to also seek to

explore the possibility of other futures for timekeeping. To do this, I will draw on work I am engaged in with colleagues at the University of Edinburgh around the idea of temporal design (Pschetz et al. 2016). That is, while the conference had a (very reasonable) emphasis on scientific and industry needs (with an odd reference to high-frequency trading), our work explores why we might want to develop forms of timekeeping that tell us about imprecise, unreliable, complicated and unpopular times. In particular, we ask what it might be like to make clocks that are not immune to, or isolated from, physical, social and political perturbations (understanding the latter term here in its broadest sense) but rather ones that respond to them.

Importantly my aim here is not to reject the interest in precision and other accompanying issues but to amplify a point that was made a number of times at the conference, namely, that different ways of calculating time are used for different purposes and that this is essential to scientific research. Building on this, my question is whether time metrologists and horologists might not also join with artists and designers in developing a greater attention to the provision of different kinds of time that address social, political and personal uses, alongside those of science and industry.

Contrasting Cartography and Horology

One reason why I am prompted to ask this question is through a comparison between horology and cartography. When it comes to making maps, it is widely accepted that there should be different kinds of maps available for different purposes. Importantly, within cartography, there is no assumption that all maps should be subsumable to a standardised, accurate, precise and homogenous space. Instead maps are more often developed through attention to the spatial problem that needs to be solved and to the optimal way of providing the information needed. Maps for using an underground transport system, for example, are very different to those used for orienteering. In fact it has been argued that a key reason why the London Underground Map has become so iconic is precisely because it does not bear a relationship to a precisely calculated space (Hadlaw 2003, p. 32).

Further, critical cartographers have argued that beyond their role in providing specific kinds of information, maps are ‘vehicles for social and political expression—of values, goals, aesthetics and status’ (Wood and Fels 1986, p. 54). The controversy of the Peters projection, which challenged the perceived Eurocentric bias of the Mercator projection, provides a good example here (see Crampton 1994). Taken together, it is clear that a map does not always need to show the ‘right space’ (e.g. according to an internationally agreed standard) in order to do its work (whether this be in simple terms or more politically complicated ones).

By contrast, however, when a clock does not show the ‘right time’, it is generally perceived that the clock has failed. Additionally, despite the work of those such as Schechner and Kevin Birth (2012), descriptions of clocks as ‘vehicles for social and political expression’ can still be met with confusion and disbelief by members of

the general public but also in certain academic circles. An important implication of this is that in comparison to maps, the idealised version of the clock—as the device that shows the 'right time'—is endowed with much less flexibility in how it can assist in solving temporal problems.

Temporal Design 'For The People'

In their introduction to critical cartography, geographers Jeremy Crampton and John Krygier argue that one of the benefits of a more politicised approach to maps is that they have been 'freed from the confines of the academic and opened up to the people' (2005, 12). In particular, they highlight artistic experiments with the form of maps as providing examples for how the theoretical critique of maps might be addressed in practice (2005, 17). Likewise, I would argue that a critically oriented approach to timekeeping could also find much inspiration in the challenges presented by artists, designers and activists to the standard form of the clock. Thus, I will turn next to a discussion of three examples from the field of temporal design which (1) challenge dominant narratives of time, (2) draw attention to alternative temporalities and (3) expose the multiplicity of time (see Pschetz et al. 2016, p. 6).

The Circadian Clock

One dominant narrative of clock time is that it consists of 24 equal hours, divided into two sets of 12. In the case of *The Circadian Clock* (2015), however, the focus is on the seasonally changing lengths of night and day (Fig. 41.1).[2] Created by Edinburgh-based designer Anaïs Moisy, the dial of the clock is engraved with shapes inspired by lace lichen (an important bioindicator of environmental health and a key component of another temporal design—*Wired Wilderness*[3]). The clock has a single hand that fully rotates once during the day and a second time at night. The rate of the rotation changes according to online information about the length of day and night in a particular location. Long summer days are indicated by a slow moving hand, while short winter days cause the hand to speed up. Thus, rather than indicating how many modern hours are left in the day, the viewer of this clock can instead read off the relative amount of day or night time left.

For Moisy, the clock particularly challenges the disconnect between clock time and the time of the local environment. However, her work also reminds the viewer of the very different ways that hours were calculated throughout European history.

[2] See http://anaismoisy.com/portfolio/conceptual-design-art/circadian_clock/.

[3] See https://climateclock.wordpress.com/residency-finalists/wired-wilderness/.

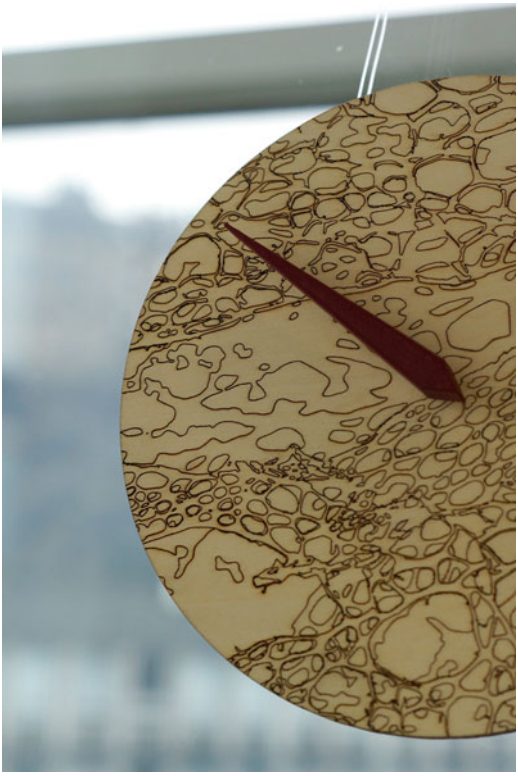


Fig. 41.1 The Circadian Clock by Anaïs Moisy ©

That is, while it might seem like common sense to work with standard hours, in the past, variable hours were used as it was more important to be able to tell the time in relation to sunup or sundown. In this way, *The Circadian Clock* challenges dominant assumptions about how clock time must operate. Moisy also demonstrates how contemporary designers can remake clocks as part of addressing temporal problems that operate according to different sets of logics.

The Coniferous Clock

Another temporal design, which in this case draws attention to alternative temporalities, is the *Coniferous Clock* (2014).[4] In its design, all that remains of the traditional clock is the circular shape (Fig. 41.2). Here, rather than mechanical moving parts and a dial, the Japanese design collective Bril has drawn on timekeeping practices linked with the production of *sake* or rice wine. Traditionally, cedar boughs are cut, shaped into balls (*sugidama*) and hung from the eaves of sake

[4] http://www.dezeen.com/2014/09/05/coniferous-clock-bril-cedar-leaves/.

Fig. 41.2 The Coniferous Clock by Bril Collective ©

breweries. When the leaves turn brown, the *sugidama* signal that the sake had matured enough to drink. In the version by Bril, a circular cedar frame is filled with fresh boughs that brown over the course of about a year. The frame can be refilled to tell time in the next year.

Resonating with Moisy's interest in bioindicators and environmental time, the *Coniferous Clock* responds to a number of factors including changes in humidity and temperature. Significantly, these are the very factors that mechanical clocks have been designed to be less responsive to. In the case of sake, however, knowing the effects of a combination of environmental conditions may be much more important. As Birth argues in relation to Pittendrigh's discovery that mosquitos coordinate their activities via humidity (rather than hours after sunset), there are always a number of environmental cycles that time can be derived from (Birth 2017). Thus, in the *Coniferous Clock*, the cultural choices around which cycles will be tracked are drawn attention to, and alternative forms of timekeeping take over the form of the clock.

Chronometers for Time Travellers

Finally, an example of a clock that emphasises the multiplicity of time is *Chronometers for Time Travellers* (2011) by artists Elaine Gan and Nik Hanselmann.[5] The project derives inspiration from Aristotle's definition of time as the measure of

[5]See http://elainegan.com/riceChrono.html.

Fig. 41.3 Elaine Gan and Nik Hanselmann, 'Chronometers for Time Travellers' installed at Digital Arts Research Center, University of California, Santa Cruz, 2011 ©

change. It consists of four containers that hold water, earth, grains and air, respectively (Fig. 41.3). On top of each container is a display that tells two types of time. The first indicates regular clock time, while the second tells the 'time' of the material. This time is calculated in reference to changes in the material which are captured through a sensor placed inside the box. Gan and Hanselmann describe their work as 'push[ing] for a rethinking of homogeneous, linear time by focusing attention on vibrant materialities and polychronic or differential tempos of change' (ibid). As I have argued elsewhere (Bastian 2012), one of the promises of the clock is that, if we follow it correctly, we will be able to coordinate with what is important to us. In *Chronometers for Time Travellers* however, the multiplicity of environmental processes is made visible, and the idea that one clock time could encompass them all is put directly into question.

Conclusion

Despite the traditional philosophical exclusion of clocks and clock time from the time of 'human existence', this paper argues for an amplification of the principle that time can be told according to a variety of reference phenomena. Drawing on a range of examples from temporal design, I have shown how the dominant form of clocks can be brought into question and alternative forms developed. I argue that by

opening up clocks 'to the people', temporal design offers another future for timekeeping, specifically one that traces a different path from more recognised efforts to advance accuracy and precision. Instead of seeking to insulate clocks from perturbations, temporal design encourages the makers of tools for time telling—whether they be scientists, designers or everyday users—to actively address political, cultural and environmental questions by liberating clocks from the assumption that they can only be correct if they tell the 'right time'.

Acknowledgements The author wishes to acknowledge the support from the UK Arts and Humanities Research Council's Connected Communities Programme during this work (grant number AH/J006637/1).

References

M. Bastian, Fatally confused: telling the time in the midst of ecological crises. Environ. Philos. **9** (1), 23–48 (2012)

K.K. Birth, *Objects of Time: How Things Shape Temporality* (Palgrave Macmillan, New York, 2012)

K.K. Birth, *Time Blind: Problems in Perceiving Other Temporalities* (Palgrave, New York, 2017)

J. Crampton, Cartography's defining moment: the Peters projection controversy, 1974–1990. Cartographica **31**(4), 16–32 (1994)

J. Hadlaw, The London underground map: imagining modern time and space. Des. Issues **19**, 25–35 (2003)

J. Hewitt Key, Calendarium, in *A Dictionary of Greek and Roman Antiquities*, ed. by W. Smith (Ed), (John Murray, London, 1875), pp. 222–233

D.C. Hoy, *The Time of Our Lives: A Critical History of Temporality* (MIT Press, Cambridge, MA, 2009)

L. Pschetz, M. Bastian, C. Speed, Temporal design: looking at time as social coordination, in DRS 2016 Design+Research+Society: Future-Focused Thinking 27–30 June 2016, Brighton, UK, 2016, Available from: http://www.drs2016.org/442/

E.G. Richards, *Mapping Time: The Calendar and Its History* (Oxford University Press, Oxford, 2000)

S. Schechner, The material culture of astronomy in daily life: sundials, science, and social change. J. Hist. Astron. **32**, 189–222 (2001)

D. Wood, J. Fels, Designs on signs/myth and meaning in maps. Cartographica **23**, 54–103 (1986)

Chapter 42
New Technologies and the Future of Timekeeping

Elisa Felicitas Arias

Abstract Time metrology and timekeeping will undergo dramatic improvements during this decade. Optical frequency standards overtake in accuracy caesium fountains by two orders of magnitude; research on time and frequency transfer will focus on techniques and methods that are adequate for the comparison of ultra-accurate clocks; time dissemination will expand; and users will become more and more demanding.

The redefinition of the SI second will be a major change in time metrology, but in parallel, new procedures will be implemented to beat the rhythm of timekeeping: better performing clock comparisons (accurate satellite time and frequency transfer, dense and permanent optic fibre networks, space-based techniques), refined algorithms for better clock characterization and new strategies for international clock comparisons.

This article will provide an inventory of the latest outcomes and extrapolate the progress in national and international timekeeping in the years to come.

Keywords SI second • Primary frequency standard • Secondary representation of the second • Time scale • Time transfer • Uncertainty

The Second: Evolution of Its Definition

The second appears in the metric system as a consequence of Gauss' work around 1832. The absolute measurements of the Earth's gravity field by Gauss promoted the establishment of the metric system based on three mechanical units: the millimetre for length, the gramme for mass and the second (of astronomical realization) for time. In 1874, the [astronomical] second constituted one of the three mechanical units of the cgs system, together with the centimetre and the gramme. After the signature of the Metre Convention in 1875, work was carried out towards the realization of the international prototypes of the metre and the kilogramme. This

E.F. Arias (✉)
Bureau International des Poids et Mesures, Sèvres, France
e-mail: farias@bipm.org

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_42

led to the creation of the MKS system: metre, kilogramme and second (the latter being realized by astronomical observations reflecting the rotation of the Earth).

The SI system was defined by the 11th General Conference on Weights and Measures (CGPM 1960) and was enlarged to its present format with seven base units following the addition of the mole by the 14th CGPM (1971) (CGPM 1972). The SI that was developed after the Second World War is a generalization of the MKSA system (A for ampere, the unit of electric current) to cover the activities of physics and technology. It is strongly related to the systems of equations between physical quantities (de Boer 1994).

The last astronomical definition of the second, as part of the SI between 1960 and 1967, was based on the orbital motion of the Earth around the Sun. The second could be realized, according to this definition, with an uncertainty of a few parts in 10^9. However, this level of accuracy could be achieved only after a decade of astronomical observations, analysis and averaging. The adoption of the atomic definition of the second in 1968, which is still in use, was an enormous step forward, not only in accuracy but also in terms of convenience. Atomic clocks existed in some institutes at the time, with relative frequency uncertainties ranging between 10^{-11} and 10^{-12}, and their industrial production had begun. The measurement of frequencies and durations was analogous to the measurements of other physical quantities, and the standard was located in the user's laboratory (Guinot 1994).

The 13th CGPM (1967) adopted an atomic definition of the second derived from the frequency of the hyperfine transition of caesium-133 (CGPM 1968). This definition must be understood as the second of proper time, for a user in the immediate vicinity of the clock and at rest with respect to it.

In 2001 the Consultative Committee for Time and Frequency (CCTF) established a working group on standard frequencies with the mission to recommend radiations as secondary representations of the second. Candidates to the list of recommended radiations are the highly accurate clocks developed in some national institutes, which include the microwave rubidium transition and a number of optical transitions realized either by optical lattices or single ions. Regular updates to the list have taken place since the creation of the working group, amounting today to eight optical radiations and one in the microwave region (List of recommended values of standard frequencies http://www.bipm.org/en/publications/mises-en-pratique/standard-frequencies.html). The best optical clocks have intrinsic uncertainties in the range 10^{-17} to 10^{-18}, two orders of magnitude better than the present caesium fountains that realize the SI second. These devices are expected to form the basis of the next definition of the second, provided that they can undergo routine comparisons without a loss of accuracy and that they can be operated over long time intervals with some flexibility.

Present Status and Resources Currently in Use

Primary Frequency Standards

The practical realization of the unit of time is understood as the definition of the unit of proper time: it applies in a small space where the caesium atoms used to realize the definition move, free of any perturbation. The unit of time is realized by caesium primary standards, most of them fountains, which are developed and operated in a small number of national institutes. The best primary frequency standards produce the SI second with uncertainties of a few parts in 10^{16} (Guéna et al. 2012; Weyers et al. 2012; Heavner et al. 2014).

A list of the caesium fountains that have contributed measurements to the BIPM for application to the computation of UTC is given in Table 42.1. The accuracy of the standards is quantified by u_B (Type B standard uncertainty), which is the combined uncertainty from systematic effects (for a complete description of uncertainties in metrology, refer to the *Guides for Expression of Uncertainty in Measurement*, GUM (2008)). This value, as provided in a measurement report, can differ from that quoted in the publication of the evaluation of the standard ($\boldsymbol{u_B}$(**Ref**)). These standards do not operate continuously, but they contribute more or less regularly to the computation of the reference timescale at the BIPM. The last column indicates the number of monthly measurement reports submitted in 2015 and the duration of the measurements, in days, ranging from 10 to 35 days. Other types of primary standards realize the SI second and contribute to UTC; they were developed prior to the caesium fountains. These "classical beam clocks" were developed in the late 1960s and have uncertainties about two orders of magnitude greater than that of fountains: today they still provide measurements on a continuous basis (Bauch 2005).

Secondary Representations of the Second

About ten optical transitions are under study at different national institutes, grouped into two categories: electromagnetically trapped single ions and multiple atoms trapped in optical lattices. The lowest published systematic uncertainties of these clocks range between parts in 10^{15}–10^{18} (Complete list of references can be found at http://www.bipm.org/en/publications/mises-en-pratique/standard-frequencies.html).

In view of the progress in the development of these ultra-accurate standards and with the aim of promoting extensive evaluations of the stability uncertainty and reproducibility of each species, the International Committee for Weights and Measures (CIPM) has maintained, since 2006, a list of frequency values and respective standard uncertainties for applications including the secondary representations of the second (CIPM 2015).

Table 42.1 Typical characteristics of the calibrations of the TAI frequency provided by the different primary (PFS) and fountain standards recommended as secondary representations of the second (SRS) reported to the BIPM in 2015

Standard	Type	Type B std. uncertainty/ 10^{-15}	u_B(Ref)/10^{-15}	Number/typical duration of comp.
IT-CsF2	PFS-Fountain	(0.17–0.30)	0.18	7/15 d–30 d
NIM5	PFS-Fountain	1.4	1.4	3/15 d–20 d
NIST-F1	PFS-Fountain	0.31	0.35	2/15 d–25 d
NIST-F2	PFS-Fountain	0.15	0.11	1/20 d
NPL-CSF2	PFS-Fountain	(0.22–0.37)	0.23	2/25 d
NPLI-CsF1	PFS-Fountain	2.82	2.5	1/10 d
PTB-CSF1	PFS-Fountain	(0.69–0.71)	1.4	8/20 d–35 d
PTB-CSF2	PFS-Fountain	(0.30–0.33)	0.41	4/10 d–30 d
SU-CsFO2	PFS-Fountain	0.25	0.50	11/25 d–35 d
SYRTE-FO1	PFS-Fountain	(0.38–0.44)	0.37	2/30 d
SYRTE-FO2	PFS-Fountain	(0.25–0.30)	0.23	11/15 d–35 d
SYRTE-FORb	SRS-Fountain	(0.28–0.32)	0.32	12/10 d–30 d

Reports of individual evaluations may be found at ftp://62.161.69.5/pub/tai/data/PFS_reports. u_B is the Type B uncertainty quantifying the accuracy of the standard at the measurement intervals, while u_B(Ref) is the value published in the standard evaluation

The uncertainties adopted by the CIPM are larger than those published in the scientific literature. These particular estimations are supported by two facts: first, the optical frequency absolute value has to be related to the present realization of the SI unit of time by the caesium frequency standards and its uncertainty, and secondly, it is preferable to retain a conservative approach considering that there are a small number of data sets, which in some cases come from a single institute.

Clocks

Industrial clocks that run continuously in contributing laboratories have an essential role in the construction of timescales. They are mostly caesium beams and hydrogen-maser clocks. The frequency provided by the industrial caesium clocks

is typically stable at the level of 1×10^{-14} over 1 day; the accuracy corresponds to an uncertainty of about five parts in ten (CIPM 2015). Hydrogen masers are characterized by a better short-term frequency stability (7×10^{-16} over 1 day).

In the construction of UTC, industrial clocks play the role of the flywheel which preserves the stability of the timescale along intervals where primary frequency standard measurements are sparse. They are also the vehicle for the broad dissemination of the reference time. There are at present about 450 industrial clocks combined in the algorithm which produces UTC; they are operated in 80 institutes which locally represent the reference timescale via UTC(k).

Industrial clocks can play different roles in a timescale. In some cases, one of the clocks in a laboratory represents itself as the local timescale, and in most cases, its output frequency can be modified with a frequency stepper to better synchronize to the reference UTC. In other cases, the timescale is computed with an algorithm which combines the readings of an ensemble of clocks, and one particular clock goes through a process of frequency steering for representing the "clock ensemble time".

These industrial clocks are used for producing more or less stable and accurate timescales depending on the operational constraints imposed by the stability and accuracy requirements. In the case of a local realization of UTC, or of a GNSS system time, the intention is to provide a "real-time" timescale. The clocks play the role of preserving the stability; the accuracy is checked by a comparison to a reference timescale (UTC).

Some institutes integrate primary frequency standards in their algorithms and produce local timescales steered to caesium fountains. This is the case for UTC (PTB), the realization of UTC at the Physikalisch-Technische Bundesanstalt (Germany) (Bauch et al. 2012), and for UTC(OP) (Rovera et al. 2016), the realization of UTC at the Paris Observatory (France) (Petit and Jiang 2008). By this procedure, the departures of these local realizations of UTC are maintained by an average of below 5 ns over periods over 1 year. Figure 42.1 shows the values of [UTC-UTC (PTB)] and [UTC-UTC(OP)] during 4 years, computed at the BIPM.

Clock Comparisons

The basic elements in the algorithm for clock ensemble timescales are the differences between readings of the clocks in the ensemble. When located in the same laboratory, clock differences are easily computed using, for example, time interval counters. Some small, known noise is contributed from these instruments to the measurements. The situation is more complex when comparing clocks in remote locations, requiring the use of techniques of time transfer which are a source of additional uncertainty.

Two techniques are used for clock comparison over relatively long distances: the time signals emitted by the two fully operational global navigation satellite systems (GPS in the USA and GLONASS in the Russian Federation) and two-way time

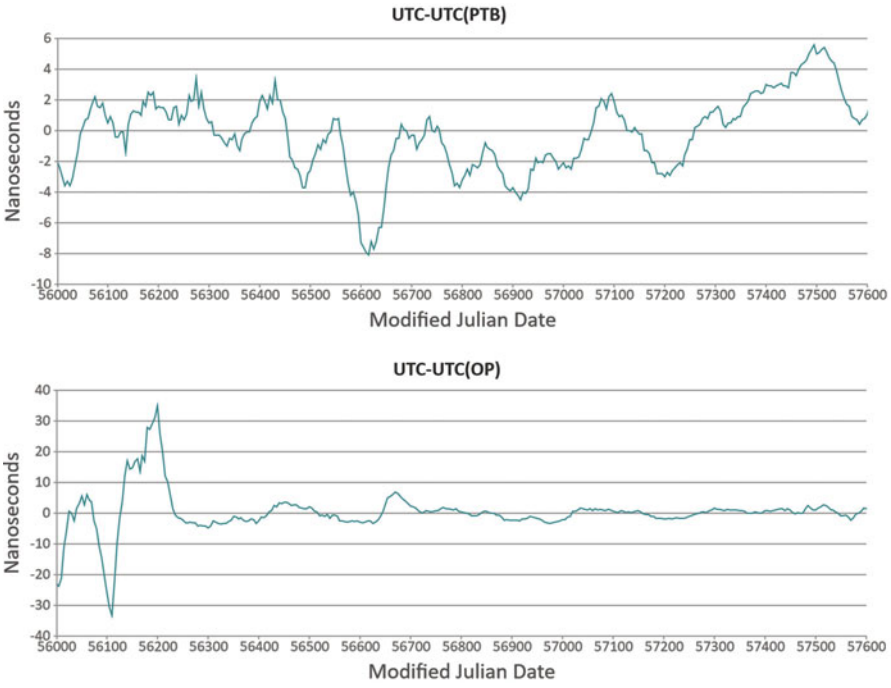


Fig. 42.1 Offsets between UTC and its local realizations steered to caesium primary standards at the PTB (*top*) and OP (*bottom*), for the interval June 2012–June 2016

transfer using (passive) telecommunication satellites (TWSTFT). Both techniques are used for regular time comparisons between institutes and are the two major techniques used in the calculation of UTC at the BIPM.

Time transfer using global navigation satellite systems (GNSS) is monodirectional; it consists of a comparison between a clock on board a satellite and a clock in a laboratory. The most frequent type of clock comparison is by common view of the same satellites at two locations; this allows the elimination of the influence of the clocks on the satellites. A more refined method has been implemented for the construction of UTC since September 2006 (Petit and Jiang 2008) making use of all the satellites in view, provided that the clocks in the satellites can be characterized. The International GNSS Service (IGS) makes available the full set of parameters necessary to correct the GNSS observations: satellite ephemerides and clocks and ionosphere total electron content maps (Products of the International GNSS Service (IGS), http://www.igs.org/products). Time comparison with sub-nanosecond statistical uncertainty with GPS is possible today by making use of the best solutions with measurements from last-generation receivers, as is the case with Precise Point Positioning (GPS PPP), a method applied on data obtained with dual frequency receivers based on the processing of the phase

and the code of the signal. This method is routinely used for precise time and frequency comparisons (Petit et al. 2011).

The TWSTFT consists of a simultaneous comparison of two remote clocks located at the two ends of a baseline (Piester et al. 2008). The equipment involved is designed for both transmitting and receiving signals. The signal goes through a geostationary telecommunication satellite which acts as a transponder. The time difference between the two remote clocks is obtained by a type of averaging of the measurements at both sides, including corrections and calibration information. It has the advantage of cancelling the effects of the ionosphere on the signal path and of dealing with a geostationary satellite. The differences between the clocks located at the two laboratories are the direct result of TWSTFT observations; it is a "link-based technique".

Reducing the uncertainty of time and frequency comparison is essential in the construction of timescales. While the frequency comparisons are only affected by the statistical uncertainty, the time comparisons are also affected by its systematic component, representing mostly the uncertainty in the measurement of the signal delay along the equipment. The characterization of the delay is known as "calibration", and it represents the largest contribution to the uncertainty of time transfer. Calibration should be regularly repeated to assess the stability of the delay; the uncertainty increases with time as a random walk process.

Calibration procedures are different in the two time transfer techniques described above. In the case of TWSTFT, the calibration is performed on a link between two laboratories. TWSTFT calibration uncertainty is typically 1 ns and is characterized by good stability, making this method very accurate. GNSS calibration is basically done on the equipment and consists of measuring the delays between the antenna and the point where the time comparison between the satellite clock and the laboratory reference is made (Jiang et al. 2011). The value for the calibration uncertainty has been confirmed to be about 2 ns in the last 2 years, thanks to regular calibration campaigns.

Growing and Future Resources

The progress in time and frequency metrology over the last decade, along with the increasing number of applications requiring stable and accurate timescales and clocks, has provoked an exponential increase in the available resources over the course of a few years. New generations of industrial atomic clocks with improved quality for use in time metrology institutes, telecoms, fundamental geodetic stations and observatories are in continuous development. Satellite and space navigation, telecommunications and defence applications require more stable and reliable clocks, with constraints on their size. Compact rubidium clocks for applications in aviation and navigation have been commercialized and have instabilities in the order of 10^{-12} between 1 and 100 s. GNSS satellites are equipped with atomic clocks that use space-qualified electronics, which are able to survive in that

particular environment, and with the highest possible performance, particularly in frequency stability and low noise for assuring predictability. Passive hydrogen masers and rubidium clocks have been designed for operating in space and can keep time within a few seconds over a period of 1–3 million years.

Optical clocks are the most accurate and stable clocks presently available, but they are still at the level of laboratory realizations. We can expect a future generation of optical clocks for industrial applications, which are currently satisfied by the performance of microwave clocks.

A new generation of microwave fountains (caesium, rubidium) are under development or already operating in time and frequency metrology laboratories; examples are the French branch of the Laboratoire national de métrologie et d'essais (LNE-SYRTE) located at the Paris Observatory, with the double caesium-rubidium fountain and the Pharao caesium clock for space applications (Abgrall et al. 2015), and the UK National Physical Laboratory (Szymaniec et al. 2016), whose caesium fountains have reached sub-10^{-16} accuracy.

Possible Redefinition of the SI Second

The development of optical frequency standards that report systematic uncertainties below those of the caesium fountain primary standards has provoked interest in considering a future redefinition of the SI second. There is, however, a compelling need for further extensive evaluations of the stability uncertainty and reproducibility of each species before a redefinition of the second can be formally considered. Among the recommended transitions (Sr, Yb, Hg optical lattice clocks and Yb^+, Al^+, Sr^+, Hg^+, Ca^+ optical lattice clocks), it is not clear which could be the candidate for the future redefinition of the second (Riehle 2015), but we can imagine that the conditions could be reached for adopting a change by 2026.

The contribution of optical clock measurements to the BIPM is another prerequisite for the redefinition of the second. Only reports from a rubidium fountain have been regularly submitted for use in the accuracy improvement of UTC. Progress in the operation of optical standards should lead to calibration intervals of the order of tens of days, which is apt for measurement submission.

Improved methods of time and frequency transfer should be operational soon to allow optical clock comparisons. Accuracies of the order of 10^{-16} are achievable for frequency transfer with GPS Precise Point Positioning with integer ambiguity resolution (Petit et al. 2015). Carrier-phase measurements with TWSTFT can reach 10^{-17} stability over long distances (Fujieda et al. 2016). Similar performance has been obtained over continental baselines using optical fibre links. A recent comparison of two strontium clocks located in Braunschweig (Germany) and Paris (France) proved that 5×10^{-17} uncertainty is possible over 1500 km (Lisdat et al. 2016). Very Long Baseline Interferometry could provide a useful tool for intercontinental comparison of very accurate standards, although its implementation in

national metrology institutes would be costly and there would be technical difficulties (Hobiger et al. 2015).

Future of Clock Comparisons for UTC

The first clock comparisons over long distances using electromagnetic signal propagation in the early 1980s used the radio navigation system Loran-C, by two-way using the telecommunications satellite Symphonie for long distances and via television signals for shorter time links. The system was supported by regular time comparisons via clock transportation. The scheme of international clock comparison consisted of time links within the continents and some long-distance intercontinental links relying on node laboratories (called "pivots"). The introduction of GPS common views for clock comparisons significantly improved the time transfer uncertainty (from microseconds to nanoseconds) and rapidly expanded in all regions. Following enhancement by TWSTFT observations in 1999, the time links became more robust and reliable. The "starlike" organization of time links was maintained until the implementation of the GPS "all-in-view" in 2006 (Petit and Jiang 2008); since then a unique laboratory has provided the pivot to all the comparisons (PTB at present).

The combination of links by GPS and GLONASS (Lewandowski et al. 2009) and by GPS and TWSTFT (Jiang and Petit 2009) is used whenever possible in the construction of UTC. Studies on the direct combination of GPS and GLONASS data have been published in 2010 (Defraigne et al. 2010, 2013), with good prospects for further improvement when the other satellite navigation systems, the European Galileo and the Chinese BeiDou, become fully operational.

Permanent links by optical fibre already exist between laboratories contributing to UTC (Śliwczyński et al. 2013); they are regularly reported to the BIPM and studied for scientific purposes. We can expect an increasing number of such links in the coming years and the possibility of including them in the time links system. This would require the design of a new scheme, somewhat similar to the previous "starlike" structure, allowing the use of optical fibres for continental time transfer.

Evolution of Timescales

Timescales will be compelled to evolve with the progress in clock and clock comparison performances. Primary and secondary frequency standards are operated in about a dozen institutes contributing to UTC; in the future, we could expect more realizations of UTC driven by these standards, requiring the development and implementation of new algorithms.

In 2012 the BIPM started an evolution towards "rapid products". The values of [*UTC-UTC*(*k*)] are published monthly in BIPM *Circular T* (BIPM *Circular T*,

monthly, http://www.bipm.org/en/bipm-services/timescales/time-ftp/Circular-T.html), with a delay of about 10 days after the end of the month of data. This allows contributing institutes to check the behaviour of their UTC(k) with respect to the reference UTC on post-real time. A rapid solution (Rapid UTC, UTCr) was developed in the Time Department and has been providing weekly predictions of the monthly results since July 2013. This solution is calculated over 4 weeks of data with an algorithm similar to that of UTC (Petit et al. 2014). With some refinement of the algorithm and the implementation of a strategy for reporting and processing primary and secondary frequency standards data, UTCr could be the test bed for the provision of *Circular T* on a shortened delay.

Conclusions

The next decade will witness dramatic changes in timekeeping, coming from the availability of better performing equipment and the implementation of more powerful analysis methods and algorithms.

Perhaps the most impressive progress has been the decreasing uncertainty of the new frequency standards. Optical clocks can provide a realization of the second that is two orders of magnitude more accurate than caesium. With this perspective, a redefinition of the SI second could be envisaged by 2026, provided that certain prerequisites are fulfilled for assessing their stability, accuracy and repeatability. Time and frequency transfer should encompass the same level of improvement, while a number of optical transitions should be regularly reported to the BIPM to improve the accuracy of UTC.

The worldwide time and frequency metrology community understood that isolated actions are not sufficient and has grouped itself into coordinated international projects that involve leading metrology institutes. Programmes supported by regional metrology organizations are addressing the milestones that need to be overcome in order to achieve the redefinition of the second.

Time transfer for the realization of UTC will benefit from multi-GNSS solutions, and, hopefully, optical fibre link data will be regularly submitted. The contributing laboratories will upgrade their equipment with multisystem receivers, and a new scheme of international time links will be put in place.

References

M. Abgrall, B. Chupin, L. De Sarlo, J. Guéna, P. Laurent, Y. Le Coq, R. Le Targat, J. Lodewyck, M. Lours, P. Rosenbusch, G.D. Rovera, S. Bize, Atomic fountains and optical clocks at SYRTE: status and perspectives. Comptes Rendus Physique **16**(5), 461–470 (2015)

A. Bauch, The PTB primary clocks CS1 and CS2. Metrologia **42**(3), S43–S54 (2005)

A. Bauch, S. Weyers, D. Piester, E. Staliuniene, W. Yang, Generation of UTC(PTB) as a fountain-clock based time scale. Metrologia **49**(3), 180–188 (2012). doi:10.1088/0026-1394/49/3/180

J. de Boer, On the history of quantity calculus and the international system. Metrologia **31**(6), 405–429 (1994)
Comptes Rendus de la 11^{e} CGPM (1960), p. 87 (1961)
Comptes Rendus de la 13^{e} CGPM (1967/68), p. 103 (1969) and *Metrologia* **4**(1), 41–45 (1968)
Comptes Rendus de la 14^{e} CGPM (1971), p. 78 (1972) and *Metrologia* **8**(1), 32–36 (1972)
CIPM Recommendation 2 (CI-2015): updates to the list of standard frequencies, http://www.bipm.org/jsp/en/CIPMRecommendations.jsp, 2015
P. Defraigne, W. Aerts, A. Harmegnies, G. Petit, et al., Advances in multi-GNSS time transfer. Proc. IFCS-EFTF **2013**, 508–512 (2013)
P. Defraigne, Q. Baire, A. Harmegnies, Time and frequency transfer combining GLONASS and GPS data, Proceedings 42nd PTTI Meeting, 2010, pp. 263–274
M. Fujieda, T. Gotoh, J. Amagai, Advanced two-way satellite frequency transfer by carrier-phase and carrier-frequency measurements. J. Phys. Conf. Ser. **723** (2016)
J. Guéna et al., Progress in atomic fountains at LNE-SYRTE. IEEE Trans. Ultr. Ferr. Freq. Contr. **59**(3), 391–410 (2012)
B. Guinot, Scales of time. Metrologia **31**(6), 431–440 (1994)
Guide to the Expression of Uncertainty in Measurement (GUM), JCGM 2008)
T.P. Heavner, E.A. Donley, F. Levi, G. Costanzo, T.E. Parker, J.H. Shirley, N. Ashby, S. Barlow, S.R. Jefferts, First accuracy evaluation of NIST-F2. Metrologia **51**(3), 174–182 (2014)
T. Hobiger, C. Rieck, R. Haas, Y. Koyama, Combining GPS and VLBI for inter-continental frequency transfer. Metrologia **52**(2), 251–261 (2015)
Z. Jiang, E.F. Arias, W. Lewandowski, G. Petit, L. Tisserand, BIPM calibration scheme for UTC time links, Proceeding of the EFTF 2011, 2011, pp. 1064–1069
Z. Jiang, G. Petit, Combination of TWSTFT and GNSS for accurate UTC time transfer. Metrologia **46**(3), 305–314 (2009)
W. Lewandowski, E.F. Arias, J. Nawrocki, P. Nogas, Use of GLONASS for international time keeping. Proc. Inst. Appl. Astron. (Ed. Nauka, St Petersburg) **20**, 358–366 (2009)
C. Lisdat et al., A clock network for geodesy and fundamental science. Nat. Commun. **7**, 12443 (2016). doi:10.1038/NCOMMS1244
G. Petit, F. Arias, A. Harmegnies, G. Panfilo, L. Tisserand, UTCr: a rapid realization of UTC. Metrologia **51**(1), 33–39 (2014)
G. Petit, A. Harmegnies, F. Mercier, F. Perosanz, S. Loyer, The time stability of PPP links for TAI, Proceeding Joint Meeting of the EFTF and IEEE FCS, 2011, pp. 1041–1045
G. Petit, Z. Jiang, GPS all in view time transfer for TAI computation. Metrologia **45**(1), 35–45 (2008). doi:10.1088/0026-1394/45/1/006
G. Petit, A. Kanj, S. Loyer, J. Delporte, F. Mercier, F. Perosanz, 1×10^{-16} frequency transfer by GPS PPP with integer ambiguity resolution. Metrologia **52**(2), 301–309 (2015)
D. Piester, A. Bauch, L. Breakiron, D. Matsakis, B. Blanzano, O. Koudelka, Time transfer with nanosecond accuracy for the realization of international atomic time. Metrologia **45**(2), 185–198 (2008). doi:10.1088/0026-1394/45/2/008
F. Riehle, Towards a redefinition of the second based on optical atomic clocks. Comptes Rendus de Physique **16**(5), 506–515., , Elsevier (2015)
G.D. Rovera, S. Bize, B. Chupin, J. Guéna, P. Laurent, P. Rosenbusch, P. Uhrich, M. Abgrall, UTC (OP) based on LNE-SYRTE atomic fountain primary frequency standards. Metrologia **53**(3), S81–S88 (2016). doi:10.1088/0026-1394/53/3/S81
Ł. Śliwczyński, P. Krehlik, A. Czubla, Ł. Buczek, M. Lipiński, Dissemination of time and RF frequency via a stabilized fibre optic link over a distance of 420 km. Metrologia **50**(2), 133–145 (2013). doi:10.1088/0026-1394/50/2/133
K. Szymaniec, S.N. Lea, K. Gibble, S.E. Park, K. Liu, P. Głowacki, NPL Cs fountain frequency standards and the quest for the ultimate accuracy. J. Phys. Conf. Ser. **723** (2016)
S. Weyers, V. Gerginov, N. Nemitz, R. Li, K. Gibble, Distributed cavity phase frequency shifts of the caesium fountain PTB-CSF2. Metrologia **49**(1), 82–87 (2012)

Chapter 43
Are Clocks Enough? Science, Philosophy, and Time

Adam Frank

Abstract It has been said that cosmology and "fundamental" physics face a crisis in that long-favored models such as string theory and the multiverse have yet to provide firm connections to observational/experimental data. Discussions of a "post-empirical science" have many in the field questioning if current approaches to quantum gravity and its applications are capable of yielding empirically coherent answers to physics' most basic questions. One aspect of these questions is attempts to rethink the role of time in physical theories. In this talk, I address the problem of time in a nonstandard way by first exploring the conflict between its scientific formulation and those that arose in twentieth-century "continental" philosophy. In particular, I will begin by addressing the famous debate between Einstein and Bergson via philosophical accounts of time that extend beyond what might be called "clock dynamics." I will then address how the view of time in phenomenology (i.e., Husserl and Heidegger) provides a quite different account of how to approach time from the standard treatments in physics. Taken as a whole, these perspectives highlight foundational questions about where and how to situate physics within the broadest account of human experience. I will end with a discussion of David Chalmers' famous hard problem and the difficulties of embedding a scientific description of time within the general problem of being as formulated in philosophy.

Keywords Time • Cosmology • Philosophy

It has been said that cosmology and "fundamental" physics face a crisis in that long-favored models such as string theory and the multiverse have yet to provide firm connections to observational/experimental data. Discussions of a "post-empirical science" have many in the field questioning if current approaches to quantum gravity and its applications are capable of yielding empirically coherent answers to physics' most basic questions. One aspect of these questions is attempts to rethink the role of time in physical theories. In this talk, I address the problem of

A. Frank (✉)
Department of Physics and Astronomy, University of Rochester, Rochester, USA
e-mail: afrank@pas.rochester.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_43

time in a nonstandard way by first exploring the conflict between its scientific formulation and those that arose in twentieth-century "continental" philosophy. In particular, I will begin by addressing the famous debate between Einstein and Bergson via philosophical accounts of time that extend beyond what might be called "clock dynamics." I will then address how the view of time in phenomenology (i.e., Husserl and Heidegger) provides a quite different account of how to approach time from the standard treatments in physics. Taken as a whole, these perspectives highlight foundational questions about where and how to situate physics within the broadest account of human experience. I will end with a discussion of David Chalmers' famous hard problem and the difficulties of embedding a scientific description of time within the general problem of being as formulated in philosophy.

Chapter 44
Time Warped: Photography, History, and Temporality

Kris Belden-Adams

Abstract Within a decade of photography's unveiling, the passenger train (1830), computer (1833), and transatlantic telegraph (1844) were introduced, followed by the invention of the telephone (1876), automobile and X-ray (1890s), cinema (1894), radio (1900–1910), airplane (1903), television (1939), Internet (1969), first personal computer (1976), and cell phone (1982). Photography and this flurry of technological advances have accelerated, and even "annihilated," the material foundations of the time/space nexus dramatically, forever redefining our already-malleable perceptions of space and time. As a consequence, time itself has been the subject of insistent theorization, speculation, and anxiety.

This poster presentation explored the relationship of photography to time and its connection to the technological forces and temporal discourses of modernity/postmodernity, which reconditioned patterns of perception. Roland Barthes, for example, wrote that the photograph has a peculiar capacity to transport the past into the present and, thus, to imply the passing of time in general. As a consequence, Barthes argued that all photographs speak of the anxious inevitability of our own death in the future. His analysis poses a challenge to all commentators on photography—what exactly is photography's relationship to time and to a shifting, increasingly complex temporal perception?

My project addresses that question by chronologically analyzing representative samples of vernacular and fine-art photography from throughout the medium's history while relating them to their technological, scientific, industrial, and cognitive contexts. The poster is available at this link: https://www.academia.edu/34088524/Poster_Presentation_Time_Warped_Photography_History_Temporality. A longer version of this study can be found at this link: https://www.academia.edu/9042293/Modern_Time_Photography_Temporality_Modernity

K. Belden-Adams (✉)
Department of Art and Art History, University of Mississippi, Oxford, MS, USA
e-mail: kkbelden@olemiss.edu

E.F. Arias et al. (eds.), *The Science of Time 2016*, Astrophysics and Space Science Proceedings 50, DOI 10.1007/978-3-319-59909-0_44

This study examines the motivations for photography's insistent struggle to reorganize time's passage, to capture or halt it, and to express time's fluctuating conditions. It argues that this struggle is both symptomatic of the unique concerns of modernity/postmodernity and is a manifestation of the photographic medium's conditional relationship to reality—a relationship that arguably has been complicated by digitalization. These trends are shaped by the medium's status as one among many technologies that redefined time and space.

Keywords Photography • Modernity • Time • Temporality • Technology

Within a decade of photography's unveiling, the passenger train (1830), computer (1833), and transatlantic telegraph (1844) were introduced, followed by the invention of the telephone (1876), automobile and X-ray (1890s), cinema (1894), radio (1900–1910), airplane (1903), television (1939), Internet (1969), first personal computer (1976), and cell phone (1982). Photography and this flurry of technological advances have accelerated, and even "annihilated," the material foundations of the time/space nexus dramatically, forever redefining our already-malleable perceptions of space and time. As a consequence, time itself has been the subject of insistent theorization, speculation, and anxiety.