THE DEVELOPMENT OF AN INTERNATIONAL ATOMIC TIME SCALE

BY

J. A. BARNES

Reprinted from the PROCEEDINGS OF THE IEEE
VOL. 55, NO. 6, JUNE, 1967
pp. 822-826

The Development of an International Atomic Time Scale

JAMES A. BARNES

Abstract—The paper reviews briefly the methods of generating atomic time and the errors inherent in the resulting scales. An atomic clock consists of an atomic frequency standard and an "integrator" to accumulate the phase of the signal. Because of noise perturbing the instantaneous frequency, an ensemble of identical atomic clocks will show a distribution of (epoch) times which is unbounded as the system evolves in time. The recognition of this problem has important consequences in national and international coordination of time scales and the construction of average atomic time scales.

Also of significance is the not completely resolved question of weighting of individual standards in the construction of average time scales. In spite of these difficulties, it is pointed out that through coordination and proper data handling, most of the advantages of astronomical time scales can be realized by atomic time scales. A statement of some of the problems facing any attempts at coordination is presented without any suggested solutions.

I. The Construction of an Atomic Time Scale

A. Introduction

When one thinks of a clock, it is customary to think of some kind of pendulum or balance wheel and a group of gears and a clock face. Each time the pendulum completes a swing, the hands of the clock are moved a precise amount. In effect, the gears and hands of the clock "count" the number of swings of the pendulum. The face of the clock, of course, is not marked off in the number of swings of the pendulum but rather in hours, minutes, and seconds.

One annoying characteristic of pendulum type clocks is that no two clocks ever keep exactly the same time. This is one reason for looking for a more stable "pendulum" for clocks. In the past, the most stable "pendulums" were found in astronomy. Here one obtains a significant advantage because only one universe exists—at least for observational purposes, and time defined by this means is available to anyone—at least in principle. Thus, one can obtain a very reliable time scale which has the property of universal accessibility. In this paper, time scale is used to refer to a conceptually distinct method of ordering events in time.

In a very real sense, the pendulum of ordinary, present-day, electric clocks is the electric current supplied by the power company. The power companies normally are careful that just the right number of swings of the pendulum occur each day, the length of the day being determined by observatories. Since all electric clocks which are powered by the same source have, in effect, the same pendulum, these clocks will neither gain nor lose time relative to each other. Indeed, they will remain close to astronomical time.

Manuscript received December 19, 1966; revised March 7, 1967.
The author is with the National Bureau of Standards, Boulder, Colo.

It has been known for some time that atoms have characteristic resonances or, in a loose sense, "characteristic vibrations." The possibility, therefore, exists of using the "vibrations of atoms" as pendulums for clocks. The study of these "vibrations" has normally been confined to the fields of microwave and optical spectroscopy. Presently, microwave resonances (vibrations) of atoms are the most precisely determined and reproducible physical phenomena that man has encountered. There is ample evidence to show that a clock which uses "vibrating atoms" as a pendulum will generate a time scale more uniform than even its astronomical counterpart.

But due to intrinsic errors in any actual clock system, one may find himself back in the position of having clocks which drift relative to other similar clocks. Of course, the rate of drift is much smaller for atomic clocks than the old pendulum clocks, but nonetheless real and important. If at all possible, one would like to gain the attribute of universal accessibility for atomic time also. This can be accomplished only by coordination between laboratories generating atomic time. Both national and international coordination are in order.

It is the purpose of this paper to review briefly the methods of constructing atomic time scales and, in doing so, to point out the limitations and difficulties facing an internationally, or, for that matter, a nationally accepted standard of atomic time (epoch). Within the literature one can find numerous papers treating time, both astronomical and atomic. It is not the purpose of this paper to review the entire field of timekeeping and show the relation of atomic time to other forms of time. For such a review, the reader is referred to the literature [1], [2]. Similarly, one may find extensive literature which covers the detailed limitations of atomic frequency standards.[1] The rather modest aim of this paper is to recognize the common difficulties of atomic frequency standards in general and develop the consequences for an international standard of atomic time. The author hopes to accomplish two things in the present treatment: first, to formulate a clear and concise statement of some of the technical (as opposed to political, personality, or traditional) problem areas to be overcome and second, to convey to individuals who are not intimately involved in the field the present, rather volatile state of affairs in atomic timekeeping.

[1] See, for example, R. E. Beehler, "A historical review of atomic frequency standards," this issue; A. O. McCoubrey, "The relative merits of atomic frequency standards," this issue; and the Special Issue on Frequency Stability, *Proc. IEEE*, vol. 54, February 1966.

It is of value in comparing time scales to consider four significant attributes of some time scales:

1) accuracy and precision,
2) reliability,
3) universal accessibility,
4) extension.

In the areas of accuracy and precision, atomic time scales have a clear advantage over their astronomical counterpart. Atomic clocks may be able to make a reasonable approach to the reliability and accessibility of astronomical clocks. The extension of time to past events (indeed, remote, past events) is a feature which atomic clocks will never possess. Their utility for future needs, however, is quite another matter. The needs of the general scientific community and, in particular, the space industries are making ever greater demands on accurate and precise timing covering longer time intervals. Often these needs cannot be met by astronomical time.

B. The Basic System

An ordinary clock consists of two basic subsystems: a periodic phenomenon (pendulum), and a counter (gears, clock face, etc.) to count the periodic events. An atomic clock differs from conventional clocks only in that the frequency of the periodic phenomenon is, in some sense, controlled by an atomic transition (atomic frequency standard) [3], [4]. Since microwave spectroscopic techniques allow frequencies to be measured with a relative precision far better than any other physical quantity, the desirability of extending this precision to the domain of time measurement has long been recognized.

It is customary to define the instantaneous (angular) frequency, $\Omega = 2\pi f$, of a signal generator by the equation

$$\Omega \equiv \frac{d\phi}{dt}, \tag{1}$$

where ϕ is the phase of the signal output and t is the time. This definition is consistent with the theory of operation of atomic frequency standards [5]–[7]. Thus, if Ω_s is the angular frequency of an atomic frequency standard and ϕ_s is the instantaneous phase, then one interprets t as atomic time. It is convenient to assume that $\Omega_s = \Omega_s(\phi_s)$, and then, the solution of (1) becomes

$$t_1 - t_0 = \int_{\phi_0}^{\phi_1} \frac{d\phi_s}{\Omega_s(\phi_s)}. \tag{2}$$

For the case where $\Omega_s = 2\pi f_s$ is constant, one may obtain from (2)

$$t_1 - t_0 = \frac{N_1 - N_0}{f_s}, \tag{3}$$

where $N_1 - N_0$ is the number of cycles (not necessarily an integer) elapsed during the interval $t_1 - t_0$ of atomic time.

It is customary to set

$$t_0 = N_0 = 0$$

at some arbitrary point in time. Several atomic scales [8]–[10] have chosen the "zero point" at zero hours, January 1, 1958 (UT2), but this is not universal among all atomic scales in existence today.

Thus, an atomic clock may consist of an atomic frequency standard and synthesizer-counter system which contains the current value of N/f_s. In practice, one normally maintains a running count of the atomic time (N/f_s) on some visual display capable of being read to the nearest second. Also a device is operated which generates a very precise electrical pulse each time the counter (N) increases its count by the numerical value of f_s (i.e., each atomic second). Fractions of one second then are determined by interpolation between the one-second ticks of the clock. For precision measurements, the usual method of interpolation is to use an electronic frequency counter operated in the time interval mode, and determine the time interval between a tick of the atomic clock and the observed event. Measurements to one nanosecond are possible by this technique and the use of "vernier methods" [11].

C. Reliability and Redundancy

In the past, reliable operation of atomic frequency standards has been a significant problem. Presently, however, commercial units with a mean time between failure (MTBF) exceeding one year are not uncommon. As with most solid-state devices, the first six to twelve months is the biggest problem, although finite atom source lifetime prevents unlimited operation without interruption.

It is true that an MTBF exceeding one year reflects significant engineering accomplishments, but this is far from comparable to the high reliability of astronomical time. The obvious solution is to introduce redundancy in the clock system. One can use several atomic clocks in the system and this should certainly be the best approach in the sense of accuracy and reliability—it is expensive, however. An alternative is to use secondary standards or crystal oscillators as "fly wheels" during down times of the primary frequency standard. A reasonable, economical compromise is probably a mixture of these two possibilities.

Suppose the synthesizer-counter subsystem of a clock system should jump a small amount and cause a discontinuity in its indicated time. It is possible that such a transient malfunction could occur with no outwardly apparent signs of malfunction of the apparatus. It is also apparent that if only two clocks are available for intercomparison, it is impossible to decide which clock suffered the transient malfunction. Thus, three clocks (not necessarily all atomic) constitute an absolute minimum for reliable operation. If one or more of these has an extended probable down time (e.g., while the atom source is replenished in an atomic device), then four or five clocks become a more workable minimum.

It should be noted here that one could assemble a large group of clocks into one system and the system MTBF calculated from the individual MTBF's might extend into geologic time intervals. This system MTBF is undoubtedly over-optimistic due to neglect of the possibilities of catastrophes or operator errors. Nonetheless, with various atomic clocks spread over the earth, it should be possible to maintain an epoch of atomic time with a reliability that could satisfy almost any future demand.

D. Propagation of Errors in an Atomic Clock

In any actual atomic frequency standard, there are always noise processes which prevent its frequency from being absolutely constant. Here it is necessary to reconcile the idea of nonconstancy with the idea of a standard. Conceptually, a standard is often defined in certain highly idealized ways. The actual physical embodiment of a standard is always less than ideal [12]. In a cesium or thallium beam device, for example, the effects of shot noise of the beam itself can be reduced by going to a high flux of atoms but the effects cannot be eliminated entirely.

Define Ω_0 now to be the "ideal" (instantaneous) frequency of an atomic frequency standard (the numerical value of Ω_0 is set by definition) and let ε represent the departure of the actual frequency Ω_s from the ideal, i.e.,

$$\Omega_s = \Omega_0 + \varepsilon.$$

Under these conditions, (2) becomes approximately

$$T \approx \frac{N}{f_0} - \frac{1}{(2\pi f_0)^2}\int_0^{\phi_1} \varepsilon(\phi)d\phi \tag{4}$$

for $|\varepsilon/f_0| \ll 1$. T is the ideal (though unobservable) time.

The most favorable class of noise which one might reasonably expect for an actual frequency standard is that ε is a band-limited white noise with zero mean. Another, entirely possible spectral type of noise, is flicker ($1/f$) noise. The sources of systematic errors and the noise sources are adequately covered in other papers (for example, Beehler et al. [12]). Defining $t_1 \equiv N/f_0$, the indicated time, (4) may be written in the form

$$t_1 \approx T + \frac{1}{2\pi f_s}\int_0^{t_1} \varepsilon_1(t)dt \tag{5}$$

where $\phi = 2\pi f_0 t$, $\varepsilon_1(t) = \varepsilon(2\pi f_0 t) = \varepsilon(\phi)$, and $|\varepsilon/f_0| \ll 1$. For $\varepsilon_1(t)$, a white noise process, the integral on the right of (5) is a "Brownian motion" or Wiener-Lévy process. A characteristic of such a process is that while its average value is zero, its excursions away from zero can be arbitrarily large as t_1 becomes large. It is easiest to imagine a large ensemble of identical clocks which were set together at $t_1 = 0$, i.e., a Dirac δ function for the initial distribution density of the clocks. As this system evolves in time, each clock will wander away from the others, and at some later time ($t_1 > 0$) there will be a spread to the distribution density of the clocks. It is a characteristic of Brownian motion that the width of this distribution increases proportionally to $\sqrt{t_1}$. If $\varepsilon_1(t)$ is other than a white noise with zero mean, the above statements do not hold. For example, if $\varepsilon_1(t)$ is a flicker ($1/f$) noise process, the uncertainty of the value of the integral grows linearly with t even though it is a nondeterministic process. It is thus important to know what types of noise predominate in actual standards and how closely they approach theoretical limitations.

The omnipresent flicker ($1/f$) noise in electronic equipment encourages one to conclude that this type of noise will be the ultimate limitation of stability of all atomic frequency standards. Within the author's experience, all atomic frequency standards *do* "flicker out" eventually when left undisturbed. The effects of occasional realignments of system parameters on the continuation of flicker noise are difficult to evaluate. It is reasonable to expect that complete alignments of the standard do destroy correlations which give rise to the flicker noise. Based on this reasoning, it is reasonable then to weight laboratory-type standards more heavily than the hermetically sealed commercial units which have been in operation for some time.

Typical values for the coefficient of the linear increase in the time uncertainty (assuming a flicker noise frequency modulation) range from 10^{-14} for hydrogen masers to a few times 10^{-13} for good cesium beams [4]. For the present state-of-the-art, the uncertainties due to systematic offsets are significantly greater than these values (10^{-11} to 10^{-12}), and thus the noise processes do not directly limit the accuracy of the instrument. The noise processes do, however, determine its precision (uniformity). One significant conclusion from these considerations may be stated: No matter how carefully systematic differences between elements of an ensemble of frequency generators are removed, the spread of times indicated by the various clocks will grow at least as $t^{\frac{1}{2}}$ and very possibly as t itself.

While these statements seem pessimistic, it is of value to recognize that Brouwer [13] determined that the random processes which affect the rotation of the earth on its axis caused the rms fluctuations in Universal Time to increase as $t^{\frac{3}{2}}$, for t greater than one year. For periods of the order of a year or less it appears that the variations in the UT2 time scale cause the rms fluctuations to increase as the first power of t (flicker noise frequency modulation). The coefficient of this linear term is about 2×10^{-9} or almost a factor of 10^4 worse than some cesium clocks.

The present means of determination of Ephemeris Time (ET) are not adequately precise to allow definitive statements about possible variations of ET [1].

It is, of course, difficult to conceive of an ensemble of solar systems to give operational meaning to some of these comments. The fact that only one solar system is used solves the problems of drifting astronomical time scales by default. The fluctuations in Universal Time are, nonetheless, observable and subject to classification by statistical techniques. A unique time scale which would be universally accessible is certainly desirable.

II. Construction of an Average Atomic Time Scale

A. Introductory Comments

As mentioned in Section I-C, there is often reason to construct average scales even within a given laboratory. In actual practice, average scales have been constructed in laboratories which, themselves, do not possess a primary atomic frequency standard. This is accomplished by using standard radio transmissions of other laboratories. This method, on the face of things, has certain advantages. By referring to a select set of primary frequency standards one hopes to accomplish two things: obtain a time scale with less bias and greater uniformity than any of the individual standards, and construct a scale which may be reproduced in any other laboratory that wishes to duplicate results. To accomplish either of the above results is difficult in practice. These difficulties arise from two sources: There are difficulties of quantitatively assessing the value of an individual standard relative to the others used in the construction of the average scale, and there exists the possibility of an inadvertent introduction of additional "Brownian motion" terms in the constructed scales.

B. Weighting Factors

There exist two possible criteria which might be considered for determining a set of weighting factors for the individual standards in establishing an average standard. A perfectly reasonable and realizable approach is to weight an individual standard inversely proportional to its mean square variation in frequency over some time interval which may be determined (at least in principle) by comparing the standard with an ensemble of other precision signal sources. With the weighting factors determined in this way, the resulting average standard should, in fact, be more uniform than any of its individual constituents (i.e., *uniformity* of rate, not necessarily *accuracy* of rate). One must consider here some of the problems of long-term stability.

An alternative is to weight the individual standards proportionally to the individual probabilities of being correct. The problem here, of course, is how to determine the accuracy of the individual standards. Does one believe the individual claims? Are the claims based on the same objective criteria? One can show that, if there is a variation in the accuracy capabilities of the individual standards and one *assumes* that they are equally reliable (equal weighting), the resulting scale is often closer to the worst in its performance than to the best. While this seems to be a significant dilemma, some very useful suggestions have come from some statistical studies [14]–[16].

Statisticians seem reasonably agreed that the simple mean (equal weighting) may, in specific situations, not be the best estimate. What is needed here is a compromise between minimum bias and safety from far-out values. The author is not aware of any laboratory that is attempting to implement the rather recent suggestions by statisticians.

In the author's opinion, the most reasonable approach is the former alternative—to base the weighting factors on the stability of the individual standards. To this end, the methods employed by Blair, Crow, and Morgan [17], [18] may prove quite useful.

C. Averages of Frequency and Time

Historically, the average time scales which have been constructed have been based on *frequency* measurements of the various standard broadcasts. Because of frequency measurement errors, it is seen that if two laboratories attempt to duplicate results in constructing average scales, relative *time* errors between the two laboratories tend to accumulate in (at best) a random walk fashion. In this situation, one has not acquired the redundancy and reliability of time measurements that is desirable.

In constructing an average time scale, two alternative data handling techniques are possible. One alternative is to treat the select set of atomic standards as simply defining frequency, as has been done in the past. In so doing, frequency measurement errors $\varepsilon_m(t)$ are introduced which are probably not correlated with errors of other laboratories observing the same set of frequency standards. Thus, each average time scale constructed on this basis has its own (independent of others) "Brownian motion." That is, even though all laboratories attempt to handle data in exactly the same way, the average scales gradually walk away from each other regardless of the fact that they are using the same set of standards and the same weightings.

The more reasonable alternative is to derive the average time scale from an average of the times (as opposed to frequencies) of the set of select time scales. This average should probably be other than the simple mean as noted above. By this technique, the measurement errors do not accumulate in an unbounded fashion as in a "Brownian motion." The problem here is obtaining comparisons of epoch for the necessary scales. While it is true that some standard broadcasts are phase-locked to their primary time scales [8], this is a fairly recent innovation and not all such broadcasts incorporate this technique. Portable clocks are an obvious, though expensive, solution [19], [20].

It should be noted that the average scale still has "Brownian motion" terms inherent in its construction. The significant point here is that the "Brownian motions" are common to all laboratories that construct the average scale and, thus, the times kept by these laboratories do *not* "walk away" from each other in an unbounded fashion. In effect, the (weighted) average epoch of the select set of atomic clocks can be considered to exist independently of its observation by some laboratory. Although there will always be some error of observation of the average, these errors of observation do not compound with the errors of subsequent observations and, hence, are not a "Brownian motion" or "random walk" type of error. By this method one has effectively recaptured the property of universal accessibility for the atomic scale.

III. Coordination of Atomic Time Scales

A. Introductory Comments

There are compelling reasons to consider an international coordination of atomic time scales to be desirable. The present paper does not pretend to solve the problems of coordination but merely to delineate and recognize some of the technical problem areas which will have to be faced by any coordination proposal. What follows is a statement of some of the technical problems which the author considers most significant.

B. Extent of Coordination

It must be decided if only one average international standard of atomic time should exist with all stations maintaining close correlation to this standard or if a more relaxed coordination should prevail. As an example, individual nations maintain their own standard of the volt and these are intercompared to define an international volt. Each country knows the relation of its volt to the international volt, but within the country the individual standards are used. An analogous system is possible with the epoch of atomic time.

In the author's opinion, a fairly close coordination is desirable. Since uniformity and precision are the most salient features of atomic time, it is inconsistent to gain coordination of atomic time scales by discrete steps in the indicated epoch of coordinating broadcast stations. It seems reasonable to maintain the broadcast time signals near the coordinated time by very small (approaching the accuracy limitations) variations in the reference frequency. The fundamental constituent time scales used in determining the one coordinated time scale would be intercompared either by portable clocks or via published values for the broadcast, coordinated time scales.

C. The Sanctity of the Individual Atomic Time Standard

Because of the problems discussed in Section II-B, most laboratories which maintain their own atomic time scales are quite reluctant to "contaminate" their scales with questionable data and techniques. Of all the problems facing coordination, this problem may well prove the most difficult. As is shown in Section II, the methods of constructing an average time scale are not closed issues. One may reason that the extreme reliability and great convenience of a closely coordinated time system should be adequate inducement to laboratories to cooperate in such an arrangement.

D. Nomenclature

As time scales have appeared throughout the world, an amazing array of different naming schemes have evolved. There exist A. 1, A. 3, TA1, NBS-A, UTC, and NBS-UA, to name a few. There is no reason why a coherent naming procedure cannot be adopted—this will probably be resolved in the near future. One could logically adopt a nomenclature which first gives the generic type of time scale (e.g., AT—atomic time) and then in parentheses the laboratory actually making the measurement [e.g., AT(NBS) would mean the atomic time scale maintained at the National Bureau of Standards]. This is quite similar to the notation used by the BIH [9].

References

[1] J. Kovalevsky, "Astronomical time," *Metrologia*, vol. 1, October 1965.

[2] G. Becker, "Von der astronomischen zur atomphysikalischen Definition der Sekunde" (in German), *PTB Mittlg.*, no. 4, pp. 315–323, and no. 5, pp. 415–419, 1966.

[3] A. O. McCoubrey, "The relative merits of atomic frequency standards," this issue.

[4] R. F. C. Vessot, H. Peters, J. Vanier, R. Beehler, D. Halford, R. Harrach, D. Allan, D. Glaze, C. Snider, J. Barnes, L. Cutler, and L. Bodily, "An intercomparison of hydrogen and cesium frequency standards," *IEEE Trans. on Instrumentation and Measurement*, vol. IM-15, pp. 165–176, December 1966.

[5] J. A. Barnes and R. C. Mockler, "The power spectrum and its importance in precise frequency measurements," *IRE Trans. on Instrumentation*, vol. I-9, pp. 149–155, September 1960.

[6] J. H. Shirley, "Some causes of resonant frequency shifts in atomic beam machines. I—Shifts due to other frequencies of excitation," *J. Appl. Phys.*, vol. 34, pp. 783–788, April 1963.

[7] J. H. Shirley, "Some causes of resonant frequency shifts in atomic beam machines. II—The effects of slow frequency modulation on the ramsey line shape," *ibid.*, pp. 789–791.

[8] J. A. Barnes, D. H. Andrews, and D. W. Allan, "The NBS-A time scale—its generation and dissemination," *IEEE Trans. on Instrumentation and Measurement*, vol. IM-14, pp. 228–232, December 1965.

[9] B. Guinot, "Temps atomique," *Bull. Horaire du Bureau Internat'l de l'Heure*, ser. J, no. 7, January–February 1965.

[10] W. Markowitz, "The system of atomic time, A. 1" *Proc. 13th Ann. Symp. on Frequency Control* (U. S. Army Signal R & D Lab., Ft. Monmouth, N. J.), p. 316, May 1959.

[11] "Practical nanosecond timer developed," *Missiles and Rockets*, pp. 34–35, May 24, 1965.

[12] R. E. Beehler, R. C. Mockler, and J. M. Richardson, "Cesium beam atomic time and frequency standards," *Metrologia*, vol. 1, pp. 114–131, July 1965.

[13] D. Brouwer, "A study of the changes in the rate of rotation of the earth," *Astron. J.*, vol. 57, pp. 125–146, September 1952.

[14] E. L. Crow, "The statistical construction of a single standard from several available standards," *IEEE Trans. on Instrumentation and Measurement*, vol. IM-13, pp. 180–185, December 1964.

[15] F. J. Anscombe and B. A. Barron, "Treatment of outliers in samples of size three," *J. Research NBS*, vol. 70B, pp. 141–147, April–June 1966.

[16] T. A. Willke, "A note on contaminated samples of size three," *J. Research NBS*, vol. 70B, pp. 149–151, April–June 1966.

[17] A. H. Morgan, E. L. Crow, and B. E. Blair, "International comparison of atomic frequency standards via VLF radio signals," *J. Research NBS*, vol. 69D, July 1965.

[18] B. E. Blair, E. L. Crow, and A. H. Morgan, "Five years of VLF worldwide comparison of atomic frequency standards," *J. Research NBS*, to be published.

[19] L. N. Bodily, "Correlating time from Europe to Asia with flying clocks," *Hewlett-Packard J.*, vol. 16, pp. 1–8, April 1965.

[20] L. N. Bodily, D. Hartke, and R. C. Hyatt, "Worldwide time synchronization, 1966," *Hewlett-Packard J.*, vol. 17, pp. 13–20, August 1966.