NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT 8400-00-84900

May 5, 1961

NBS REPORT 6766

ATOMIC FREQUENCY STANDARDS AT THE NATIONAL BUREAU OF STANDARDS

J. M. Richardson, R. E. Beehler, R. C. Mockler, and R. L. Fey
National Bureau of Standards
Boulder, Colorado

A report submitted to the Comité Consultatif pour la Definition de la Secônde auprés du Comité International des Poids et Mesures

for its sessions of April 11-12, 1961

NBS

U. S. DEPARTMENT OF COMMERCE
NATIONAL BUREAU OF STANDARDS
BOULDER LABORATORIES
Boulder, Colorado

ATOMIC FREQUENCY STANDARDS AT THE NATIONAL BUREAU OF STANDARDS

J. M. Richardson, R. E. Beehler, R. C. Mockler and R. L. Fey
National Bureau of Standards
Boulder, Colorado

Abstract

Many laboratories in various countries have now shown that certain atomic resonance frequencies have such desirable properties that they may be considered suitable for a definition of the unit of time. Details in the performance of actual atomic resonance devices are given. Also, estimates are made of how closely a given device approaches the idealized resonance frequency.

Recently, two cesium beam frequency standards of independent design and construction have been completed, evaluated, and compared at Boulder. The main characteristics and performance details of these standards are analyzed and listed.

An analysis is given of comparisons of the U. S. Frequency Standard, by standard frequency broadcasts, with other frequency standards in the world.

ATOMIC FREQUENCY STANDARDS AT THE NATIONAL BUREAU OF STANDARDS

J. M. Richardson, R. E. Beehler, R. C. Mockler and R. L. Fey
National Bureau of Standards
Boulder, Colorado

Many laboratories in various countries have now shown that certain atomic resonance frequencies have such desirable properties that they may be considered suitable for a definition of the unit of time. In such a case, it is essential to know in detail the performance of actual atomic resonance devices and to estimate how closely a given device approaches the idealized resonance frequency, ν_o. This degree may be spoken of as the accuracy of the device with respect to ν_o, and it will indicate how well a practical device can be made to realize the idealized definition.

Recently, two cesium beam frequency standards of independent design and construction have been completed, evaluated, and compared at Boulder[1,2]. The main constructional characteristics of the devices are listed in Table I.

Several sources may limit the accuracy (in the above sense) of an atomic frequency standard. These include the uniform magnetic C field, phase differences between the two oscillating electromagnetic fields, the spectral purity of the radiation exciting the atomic resonance, and the presence of other neighboring atomic transitions. For the case of the NBS standards possible frequency errors arising from these and other sources have been carefully investigated. The results are also given in the accompanying table. In those cases where a frequency shift is measurable so that a correction may be applied to the measurement (extrapolation to zero magnetic field, for example) the shift is not considered to be a contribution to the inaccuracy of the device; however, any uncertainty in the amount of shift <u>is</u> considered an inaccuracy.

The uniform C fields are normally adjusted to be within the 0.035 to 0.050 oersted range. They are produced by passing a current through a rigid trough-like conductor in NBS I and a rectangular array of four parallel wires in NBS II. Shielding against stray fields is accomplished with a single mu-metal cylindrical shield and a triple layer mu-metal and soft iron shield assembly, respectively. The measured field non-uniformity is less than ± .001 oersted. The frequency correction to be applied for the C field used in a measurement is determined by measuring the field in terms of certain strongly field-sensitive transitions, such as the $(F = 4, m_F = 1) \leftrightarrow (F = 3, m_F = 1)$ or the $(F = 4, m_F = 1) \leftrightarrow (F = 3, m_F = 0)$ transitions. Measurements for confirmation are also made using the $(\Delta F = 0, \Delta m_F = \pm 1)$ low-frequency transitions and a high-sensitivity oerstedmeter. In the field measurements the greatest reliance is placed upon the microwave measurements, since the spectral lines at low frequencies are more subject to distortion and power shifts. The small differences of 0.002 oersted for NBS I and 0.001 oersted for NBS II observed among the various methods of measurement are taken as the uncertainties in the field determinations, producing the corresponding uncertainties in frequency listed in the table.

Another correction to the frequency measurements is necessary because of existing phase differences between the separated oscillating fields caused by non-uniform absorption of microwave power over the surfaces of the cavities. The effect is observed physically by rotating the resonant cavity structure by 180° and looking for a frequency shift. With NBS II a small relative shift of 4×10^{-12} was recently observed, resulting in a correction of one-half this amount or 2×10^{-12} to the measured frequency. A somewhat larger shift of 1.6×10^{-11} was observed with NBS I, possibly as a result of visible imperfections in the electroformed cavity walls.

A third source of possible errors is the nearness of neighboring transitions in the atomic spectrum. Significant frequency shifts may result in measurements made at low C fields. For measurements made with NBS II at a field of 0.020 oersted, for instance, a systematic shift of 3.7×10^{-11}

was detected. It has been found possible to eliminate this error by operating at sufficiently high fields (0.047 oersted for NBS II).

A final factor which has been observed to cause large frequency shifts under some conditions is the spectral purity of the radiation exiting the atomic transition.[3] Shifts of 32×10^{-10} have been observed by exciting the resonance with a signal containing unsymmetrically-placed sidebands about the carrier frequency. These effects can also be eliminated by utilizing an oscillator phase-locked to the oscillator to be measured with time constants suitably chosen to make use of the long-term stability of the original oscillator and the short-term stability of the phase-locked oscillator. Additional uncertainties due to applied power level and resonant cavity detuning have been found to be negligible under normal operating conditions.

The uncertainty in approaching ν_o because of the above effects is shown in the table. To obtain a limit of error for each machine we must also add the random uncertainty associated with the measurements themselves, as limited by the stability of the flywheel oscillator. One observation of frequency on either device requires about 15 seconds. The standard deviation of one such measurement is normally about 1×10^{-11}, while the standard deviation of the mean for 15 - 20 such measurements made over about 10 minutes is 2×10^{-12}. This then is taken as the precision of the measurement of frequency and of frequency difference between the two machines.

If it is assumed that the uncertainties are interdependent, then the estimate of the total error in absolute frequency measurement is obtained by adding all the uncertainties. Estimated in this way, NBS I has an accuracy of $\pm 0.8 \times 10^{-11}$ and NBS II has an accuracy of $\pm 1.1 \times 10^{-11}$. The difference frequency between the two should fall within $\pm 1.9 \times 10^{-11}$ (standard deviation).

If on the other hand the various uncertainties in frequency are considered independent, they should properly be squared, added and the square root extracted for the estimate of the limit of error. With this assumption, the estimated error is $\pm 0.5 \times 10^{-11}$ for NBS I and $\pm 0.7 \times 10^{-11}$ for NBS II. The difference frequency should fall within $\pm 0.8 \times 10^{-11}$ (standard deviation). The measured average frequency difference is:

$$(\text{NBS II} - \text{NBS I})_{av} = -(1.6 \pm 0.3) \times 10^{-11}.$$

The estimate of error and the measurements appear then reasonably consistant. We state the accuracy with which either machine approaches ν_o to be about $\pm 2 \times 10^{-11}$ and the precision to be $\pm 2 \times 10^{-12}$.

Frequency comparisons between NBS II and NBS I have been carried on for over one year. Since the NBS standards are not used with servo oscillators locked to the atomic resonance, the machines are compared by noting the difference between successive measurements on both machines of a very stable quartz oscillator or an ammonia maser. The best value for the difference between the two machines at present is that given above. Analysis of comparison data from February to March 1960 and some recent data of February 1961 shows that under similar conditions of the C fields during the two periods the relative frequency difference between NBS II and NBS I has not changed by more than 2×10^{-12} in one year.

Regular comparisons of Atomichron R 106 at Boulder with NBS I and NBS II have also been made over a period of nearly two years. During this time the beam tube in the Atomichron R has been replaced twice. The first replacement in October 1959 resulted in a relative Atomichron R frequency shift of 3×10^{-10}; the second replacement in May 1960 produced a frequency shift of only 5×10^{-11}. During the last 15 months when the Atomichron R has been measured on almost a daily basis, the mean frequency difference observed was 0.9×10^{-10} with a maximum range of 3.2×10^{-10} for the approximately 400 comparisons. The mean deviation was 4×10^{-11} and the standard deviation was 6×10^{-11}.

Systems utilizing the Cs resonance to stabilize quartz oscillators have not as yet been developed to an extent which makes full use of the stability exhibited by beam machines functioning as passive resonators.

The United States Frequency Standard (NBS I or NBS II as described above) is made available for comparison with other frequency standards in the world through the National Bureau of Standards Radio Stations WWVB, 60 kc and WWVL, 20 kc both located near Boulder, Colorado. The carrier frequency of WWVB is derived by direct synthesis from the frequency of an Atomichron **R**. That of WWVL is derived from a crystal oscillator, and its transmitted frequency is received and compared to that of the Atomichrons **R** each day by means of a 24 hour average. The Atomichron **R** frequency is in turn compared with that of the NBS atomic standard once each day with an averaging time of about 10 minutes. Other received LF and VLF signals are measured in the same way. This procedure unavoidably degrades the stability available from NBS I or NBS II in its use for assigning values to the transmitted or received carrier frequencies, as a consequence of the observed variability of the Atomichron R mentioned above.

The daily frequencies of WWVB and WWVL as received and published by other laboratories with atomic standards are used to obtain monthly means which are customarily plotted to give an indication of the agreement between the NBS standard and those of the other laboratories. By using multiple comparisons when no direct link is available between the Boulder Laboratories and another standard, relations between all standards may be displayed. Based on these plots, a few general statements may be made concerning agreement among standards: Differences between the two most divergent of the eight standards, NBS, NRL, Naval Observatory, Cruft, NRC, NPL, CNET, and Neuchatel were as great as 8 parts in 10^{10} and as small as 3 parts in 10^{10} during the months September, 1960 through February, 1961. A more typical mean difference between stations was perhaps 1 or 2 parts in 10^{10} with fluctuations from month to month of the same order. There were about as many standards above Boulder Laboratories in frequency as below, with some crossover occurring as time progressed.

Sufficient data have now accumulated, however, from a number of monitoring stations to permit a detailed statistical study which will give more quantitative results than that just presented. Therefore, from these data the period of September, 1960 through February, 1961 was again selected and subjected to a standard two way analysis of variance[4].

The model upon which this analysis was based involves the assumption that there are two factors which may systematically affect the result of a comparison of frequencies between pairs of standards by means of radio propagation. These are:

(1) Systematic differences which may exist among the frequencies of the various standards.

(2) Systematic differences which may be observed in received frequencies from day to day due to fluctuations of the transmitter frequency and fluctuations of the propagation medium.

In addition residual fluctuations not assignable to either of these causes are considered to be other random errors of measurement such as day to day variations in radio propagation not common to all the paths to receivers; day to day variations of the individual monitoring atomic standards; and other day to day variations of the receivers.

The data from one month's measurement of the carrier frequency of one station by as many laboratories as possible has been chosen as a convenient unit for one statistical test; the null hypothesis under test is then that the means of the populations from which the observations were drawn are all equal assuming that averaging is done in either of two ways:

(a) over all of the standards on a given day or
(b) over all of the days for a given standard.

If the hypothesis proves false by the test, that is, if statisically significant differences exist among standards and among different days, the analysis also provides a measure of the size of these differences as well as a measure of the random errors of observation.

The analysis was carried out for each of the transmitters, GBR, NBA, WWVL, and WWVB, using data from all laboratories which monitored these stations. For every station there were some days of each month for which data were unavailable from one or more of the participating laboratories, typically permitting the use of about 12 to 20 days data for one month.

The results of the analysis were generally that statistically significant differences do exist among standards and among days, and are presented in Table II as standard deviations assignable to the indicated causes. All entries are statistically significant at the 5% level by the Fisher F test except those indicated by an asterisk. The laboratories which monitored each transmitter are listed at the bottom of the column for that transmitter. The entries in Table II should be understood in the following context: If, in a given month, a single one-day observation of frequency of a particular transmitter were made by a randomly chosen receiving station, then the total standard deviation, σ_T, of this observation is the square root of the sum of the squares of the three standard deviations displayed in Table II, i. e.,

$$\sigma_T = (\sigma_S^2 + \sigma_D^2 + \sigma_R^2)^{1/2}$$

where σ_S, σ_D, and σ_R refer respectively to the standard deviations due to differences among standards, differences among days, and residual differences.

The results for each indicated cause of variation may be considered briefly.

(1) Differences among standards may be expected to be in the neighborhood of 1 to 2 parts in 10^{10}.

(2) Differences among days are seen to be somewhat larger for GBR. This may be expected since the average path length was greatest for the group of laboratories which monitored GBR.

(3) The entries for residual differences are generally similar, indicating that the various measuring techniques have comparable precision. Agreement from month to month for GBR and also for WWVL indicate that no large unknown factors are affecting the measurement.

Data available to us at this writing for NBA are insufficient to make comparisons, and those for WWVB are unreliable because of its low transmitted power.

Although the cesium beam is the device which has enjoyed the highest degree of refinement, the most widespread use, and the widest intercomparison, nevertheless other devices based on other atoms or molecules must be considered. These are the NH_3 beam maser, the Rb gas cell resonator, the atomic hydrogen maser, and the Tl beam. Extensive experience has been had with $N^{14}H_3$ maser[5], and this experience leads to the conclusion that the $N^{14}H_3$ maser should not be considered a serious competitor to the Cs beam device. Although one can derive an oscillatory signal of extreme spectral purity (a few parts in 10^{12}, better than any other r-f source) from the maser, it suffers from frequency shifts of the order of several parts in 10^9 dependent on the details of operation, such as beam flux, focuser voltage, cavity tuning, choise of isotope, and incompletely eliminated Doppler effect. It can be said that the molecules in such a maser are not so well isolated as the atoms in the Cs beam, and so are subject to greater perturbation. The maser conditions of operation must be specified more carefully than the Cs conditions in order to achieve comparable reproducibility.

Gas cell frequency standards have been investigated in a number of laboratories[6]. Rubidium appears to be the most suitable element to work with, and it has accordingly received the most attention. The hyperfine frequency of rubidium in a gas cell can be determined to very high precision using optical pumping techniques. However, the frequency can be affected by the buffer gas. Since background contaminants will affect the frequency, a gas cell frequency standard can be expected to drift as the buffer gas changes its density or composition. Cells have been produced so that this aging is less than one part in 10^{11} per month, and it is expected that it may be possible to make the aging considerably less than that. Sensitivity to temperature can be made small by a judicious choice of the buffer gas composition. Cells can be made reproducibly to the same frequency with errors less than one part in 10^{10}.

The gas cell frequency standards have been found to be somewhat sensitive to the intensity of the pumping radiation. The energy levels can be shifted by a distribution of incident light which is asymmetric with respect to the absorption lines. This radiation causes virtual transitions to the excited state.

The resultant change in the energy is different for the different hyperfine sublevels, and appears as a change in the hyperfine resonance frequency. It is possible to arrange conditions such that the shift in frequency due to the optical radiation is relatively small. The long term shift in frequency because of this effect, may be as large as one part in 10^{10}.

The atomic hydrogen maser[7] holds great promise, possibly to precisions of parts in 10^{13}. Since it is newly developed, we presently lack extensive and widespread knowledge of its performance characteristics. It is treated in a separate report submitted to this Committee.

The resonance of $T\ell^{205}$ at 21,311 Mhz has been suggested as worth investigating because of its higher frequency, greater mass, and lower sensitivity to magnetic field. So far a $T\ell$ beam has been detected at the National Bureau of Standards. Further work is in progress.

From the foregoing discussion, we can conclude that the present degree to which a carefully constructed and tested Cs standard can approach the idealized Cs resonance is of the order of a few parts in 10^{11}. This is without regard to special or arbitrary prescriptions of design and operation, but only with regard to knowledge of the pertinent parameters to sufficient accuracy. Comparison of different Cs devices gives agreement within a few parts in 10^{10}. The precision of intercontinental comparison by VLF radio can be made somewhat better than 1 part in 10^{10}. In passing, it is interesting to note that the precision of intercontinental frequency comparison is about equal to the precision of the present standards; if superior standards at the level of 10^{-13} precision, say, are developed, it may be difficult to compare them. Research and development will, of course, inevitably continue at a rapid pace so that standards superior to the Cs standard may reasonably be expected in the long term future. However, the emergence of such a superior standard will merely reduce to the simple problem of comparing it with existing Cs standards rather than a new comparison with the (Ephemeris) second, so that there will be no bar to easy future improvement in the unit of time.

We therefore believe that it is both feasible and advantageous to recommend definition of the second in terms of the $(F = 3, m_F = 0) \leftrightarrow (F = 4, m_F = 0)$ transition of Cs^{133} on the basis of the best information available before 1966 as to the transition frequency in terms of the present second.

Acknowledgment

This report represents the work of many people. The authors wish to acknowledge the contributions of A. H. Morgan, D. H. Andrews and P. A. Simpson who are responsible for the standard frequency broadcasts of the National Bureau of Standards and for the comparison of the U. S. Frequency Standard with other frequency standards in the world via radio broadcasts.

The statistical techniques used were suggested by Dr. E. L. Crow.

Dr. E. C. Beaty contributed the description of the Rb cell.

Finally we wish to thank the other members of the Atomic Frequency Standards Section, particularly Mr. Charles Snider, for their work on the United States Frequency Standard.

References

1. R. E. BEEHLER, R. C. MOCKLER and C. S. SNIDER, "A comparison of atomic beam frequency standards", Nature, 187, 681, (1960).

2. R. C. MOCKLER, R. E. BEEHLER and C. S. SNIDER, "Atomic beam frequency standards", Trans. IRE, I-9, 120 (1960).

3. J. A. BARNES and R. C. MOCKLER, "The power spectrum and its importance in precise frequency measurements", Trans. IRE, I-9, 149 (1960).

4. E. L. CROW, F. A. DAVIS and M. W. MAXFIELD, Statistics Manual, Dover Publications, Inc. (1960) Chapter 5.

5. R. C. MOCKLER, J. BARNES, R. BEEHLER, H. SALAZAR and L. FEY, "The ammonia maser as an atomic frequency and time standard", Trans. IRE, I-7, 201 (1958).

6. R. J. CARPENTER, et al, "A prototype rubidium vapor frequency standard", Trans. IRE, I-9, 132 (1960).
W. E. BELL and A. L. BLOOM, "Optically detected field independent transitions in sodium vapor", Phys. Rev. 109, 219 (1958).
J. M. ANDRES, D. J. FARMER and G. T. INOUYE, "Design studies for a rubidium gas frequency standard", Trans. IRE, Mil-3, 178 (1959).
P. L. BENDER, E. C. BEATY and A. R. CHI, "Optical detection of narrow Rb^{87} hyperfine absorption lines", Phys. Rev. Letter 1, 311 (1958).
M. ARDITI and T. R. CARVER, "Frequency shifts of microwave resonances in gas cell using optical pumping", Proceedings of the 14th Annual Symposium on Frequency Control, 329 (May 1960).

7. H. M. GOLDENBERG, D. KLEPPNER and N. F. RAMSEY, "Atomic hydrogen maser", Phys. Rev. Letters, 5, 361 (1960).

TABLE I. CHARACTERISTICS OF NBS CESIUM BEAM FREQUENCY STANDARDS

	NBS I	NBS II
Characteristics		
Beam dimensions	.008 × .254 cm	.038 × .475 cm
Separation of oscillating fields	55 cm	164 cm
Line width	300 hz	120 hz
Magnetic shields	single	triple
Cavity Q	5000	5000
Corrections to Frequency		
C field	5×10^{-11}	9×10^{-11}
Phase shift	0.8×10^{-11}	0.2×10^{-11}
Power spectrum	0	0
Neighboring resonances	0	0
Power level, cavity detuning	0	0
Uncertainties in Frequency		
C field	$\pm 0.4 \times 10^{-11}$	$\pm 0.5 \times 10^{-11}$
Phase shift	$\pm 0.2 \times 10^{-11}$	$\pm 0.4 \times 10^{-11}$
Power spectrum	0	0
Neighboring resonances	0	0
Power level, cavity detuning	0	0
Measurement precision	$\pm 0.2 \times 10^{-11}$	$\pm 0.2 \times 10^{-11}$
Total	$\pm 0.8 \times 10^{-11}$	$\pm 1.1 \times 10^{-11}$

Measured difference, $(\text{NBS II} - \text{NBS I})_{av}$ - - - - - - - - - - - - -1.6×10^{-11}

TABLE II. RESULTS OF TWO WAY ANALYSIS OF VARIANCE

(Entries are in units of 10^{-10}.)

		Transmitting Station GBR	NBA	WWVL	WWVB
Freq.		16kc	18kc	20kc	60kc
Rad. Power		40kw	40kw	15w.	1.5w
(a) Standard deviation,	Sept.		1.08	.43*	3.7
σ_S, associated with	Oct.	2.02	.98*	.58	3.7
mean differences among	Nov.	1.79		.70	10.2
standards at the indicated	Dec.	2.16		1.76	1.1
monitoring laboratories	Jan.	1.26		.99	1.1*
	Feb.	2.10		1.06	
weighted average		1.94	1.03	1.10	
(b) Standard deviation,	Sept.		.84	.93*	.00*
σ_D associated with	Oct.	1.82	.00*	.78	.66*
mean differences among	Nov.	5.27		2.70	1.05*
days as measured by the	Dec.	4.16		2.09	.78
same monitoring labor-	Jan.	2.48		2.60	1.08
atories as above	Feb.	1.05		1.95	
weighted average		3.57	.55	1.96	
(c) Standard deviation,	Sept.		.33	2.10	2.9
σ_R associated with	Oct.	1.08	3.02	.35	.9
residuals as measured	Nov.	1.61		1.45	9.5
by the same monitoring	Dec.	1.88		1.15	1.4
laboratories as above	Jan.	1.16		1.62	2.8
	Feb.	.90		1.71	
weighted average		1.15		1.18	
Laboratories monitoring each transmitting station		NBS NRC NPL CNET Cruft	Nav. Obs. NRL NPL CNET Cruft Neuchatel	NBS NRL Cruft	NBS NRL NRC Cruft

* Not statistically significant
(All other entries significant at the 5% probability level)