Accuracy evaluation of NIST-F1

S. R. Jefferts, J. Shirley, T. E. Parker, T. P. Heavner, D. M. Meekhof, C. Nelson, F. Levi, G. Costanzo, A. De Marchi, R. Drullinger, L. Hollberg, W. D. Lee and F. L. Walls

Abstract. The evaluation procedure of a new laser-cooled caesium fountain primary frequency standard developed at the National Institute of Standards and Technology (NIST) is described. The new standard, NIST-F1, is described in some detail and typical operational parameters are discussed. Systematic frequency biases for which corrections are made – second-order Zeeman shift, black-body radiation shift, gravitational red shift and spin-exchange shift – are discussed in detail. Numerous other frequency shifts are evaluated, but are so small in this type of standard that corrections are not made for their effects. We also discuss comparisons of this standard both with local frequency standards and with standards at other national laboratories.

1. Introduction

We present the evaluation procedure for NIST-F1, a laser-cooled caesium fountain primary frequency standard operated by the NIST Time and Frequency Division in Boulder, Colorado, USA. NIST-F1 has been operated as a frequency standard since November 1998 and has undergone eleven frequency evaluations (five reported to the Bureau International des Poids et Mesures (BIPM), six for internal use) in the interval between November 1998 and August 2001. The results of these evaluations are shown in Section 7 (Figure 11) of this document. Five of the eleven evaluations, with Type A (statistical) uncertainties less than 1.5×10^{-15} and Type B (systematic) uncertainties less than 1.4×10^{-15}, have been reported to the BIPM for inclusion in TAI (Temps Atomique International/International Atomic Time). The first two reported evaluations used measurements of atom density and published coefficients to estimate the spin-exchange shift. The third and subsequent evaluations were made using a range of atom densities and extrapolating the observed frequency to zero atom density.

The paper is divided into seven sections. Section 2 is a description of the standard. This is followed in Section 3 by a discussion of biases for which corrections are made. Section 4 discusses biases for which the standard is uncorrected. Section 5 is a brief discussion of the short-term stability of the standard, while Section 6 gives an overview of the evaluation procedure. Finally, in Section 7 there is a discussion of the local timescale with which the standard is compared, the cumulative results and the transfer process.

2. Description of NIST-F1

2.1 Overview

NIST-F1 is designed as a primary frequency standard. Accuracy and long-term stability are the ultimate goals. Short-term stability of the reference, while important for operational reasons, is a secondary concern.

The fountain described here, like the Laboratoire Primaire du Temps et des Fréquences (LPTF) fountain [1], uses a (0,0,1) geometry for the laser cooling and launching operation in which four of the six laser beams are in the horizontal plane while two are vertical. Figure 1 illustrates the vacuum chamber and associated structure on NIST-F1.

The two vertical laser beams are used to launch the atoms. A sample of about 6×10^6 Cs atoms is first cooled in an optical molasses with six laser beams at 852 nm. The sample of cold (approximately 1.3 μK) atoms is launched vertically upwards from the optical-molasses region with an initial velocity which is typically about 4 m/s. The atom sample

S. R. Jefferts, J. Shirley, T. E. Parker, T. P. Heavner, D. M. Meekhof, C. Nelson, F. Levi,[2] G. Costanzo,[3] A. De Marchi,[3] R. Drullinger, L. Hollberg, W. D. Lee and F. L. Walls: Time and Frequency Division, National Institute of Standards and Technology (NIST), 325 Broadway, Boulder, CO 80303, USA.
[2]Permanent address: Istituto Elettrotecnico Nazionale "G. Ferraris", Str. Delle Cacce 91, 10135 Turin, Italy.
[3]Permanent address: Politecnico di Torino, C.so Duca degli Abruzzi, 24, 10129 Turin, Italy.

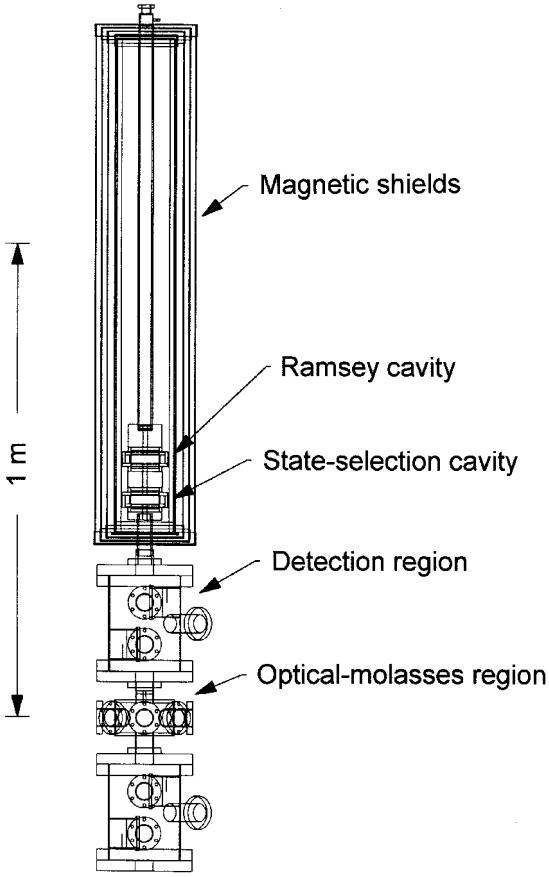

Figure 1. Mechanical drawing of the NIST-F1 physics package. The relevant parts of the assembly are shown along with a scale factor.

drifts upwards through the detection region (which is turned off when the atoms travel up) and enters the magnetically shielded C-field region. The magnetic field (C-field) that provides a quantization axis in the fountain is small (in comparison with thermal beam standards), about 10^{-7} T (10^{-3} G). After entering the C-field region the atoms enter the state-selection cavity which is used to select atoms in the "clock" state ($|F, m_f\rangle = |F, 0\rangle$). The state selection cavity is a TE_{011} cavity which transfers $|4, 0\rangle$ atoms to the $|3, 0\rangle$ state. Essentially all (> 99.9 %) of the atoms remaining in the $F=4$ state are then removed from the measurement sample with an optical pulse from the vertical laser beams. The atoms next encounter a TE_{011} microwave cavity where microwave excitation is performed. After having passed the excitation cavity on the way up, the atoms (now in a superposition state) continue to decelerate under the influence of gravity. Eventually the atoms reach apogee and begin to fall. Some fraction (roughly 10%) of the atoms (determined primarily by the atom temperature and toss height) re-enter the excitation cavity. The time separation between the two passages through the excitation cavity has the same effect on the atoms as Ramsey's (spatially) separated oscillatory-fields method. The atoms continue to fall, eventually leaving the C-field and entering the detection region (which has been turned on by this time) where the relative atom populations in the $F=3$ and $F=4$ hyperfine levels are measured. The measured populations of the $F=3$ and $F=4$ levels are combined to generate an error signal used to steer the microwave frequency synthesizer on to the frequency of the atomic transition. This constitutes one cycle of the pulsed fountain operation.

2.2 Physics package

2.2.1 Optical-molasses region

While the fountain does have a magneto-optical trap (MOT), this feature has not been used during any of the measurements reported here. In general linear-perpendicular-linear (lin⊥lin) molasses is used as the method of gathering the atom sample. Table 1 gives typical parameters for the atomic source region.

Table 1. Typical parameters for operation of the caesium fountain as a primary frequency standard.

Parameter	Symbol	Value
Gather atoms into molasses	T_{gather}	0.40 s
Total atoms gathered and tossed	N_0	6×10^6
Atoms in $m=0$	$N_{0,m=0}$	7×10^5
Atoms detected	N_{detected}	6×10^4
Atom temperature (after launch)	T_{atoms}	1.3 μK
Typical launch velocity	V_{launch}	4.25 m/s
Ramsey time	T_{Ramsey}	0.56 s
Launch height above Ramsey cavity		0.38 m
Optical intensity (horizontal beams)		7 mW/cm^2
Beam diameter (horizontal beams)		2.5 cm
Optical intensity (vertical beams)		6 mW/cm^2
Beam diameter (vertical beams)		1 cm

2.2.2 C-field region

The C-field region is magnetically shielded using four layers of 79% nickel-iron alloy. The shields are concentric cylinders with end-caps. The upper end of the shield cylinders have the end-caps welded in place, while the lower end-caps are clamped in place. The shields have holes in the end-caps for the vacuum structure. The holes are approximately 3.5 cm in diameter in the upper end-caps and 4.2 cm in diameter in the lower end-caps. Inside the innermost shield is the C-field solenoid which is 1.04 m long, 0.12 m in diameter, and has a total of 410 turns.

Figure 2 shows the measured magnetic field inside the C-field region. This field map was generated by measuring the transition frequency of the $|3, 0\rangle \rightarrow |3, m \neq 0\rangle$ magnetically sensitive transition using a low-frequency excitation coil transverse to the flight path. The effect of field leakage from the holes in the end-caps can be clearly seen (near 80 cm in Figure 2), as can the effect of the highly permeable vacuum feedthroughs on the microwave cavities. While

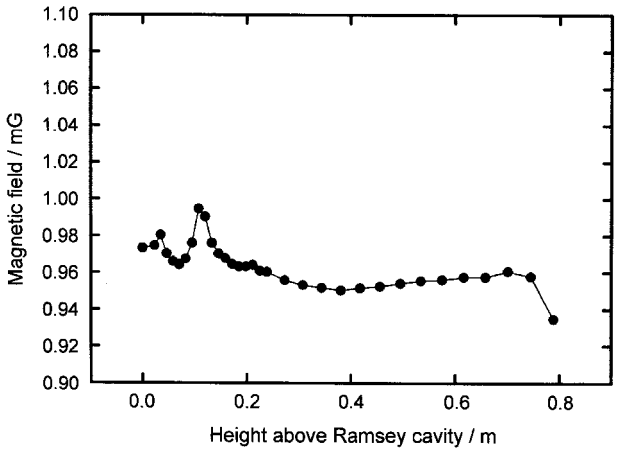

Figure 2. Map of the magnetic field in and above the Ramsey cavity. The origin of the x axis is the centre of the Ramsey cavity. The vacuum feedthroughs on the Ramsey cavity are quite permeable, and the resulting distortions to the magnetic field are "corrected" by shim coils placed near the microwave cavity. This is the cause of the field distortions shown.

the magnetic field is not as homogeneous as might be desired, the effects of the inhomogeneity are small, as discussed in Section 3.2.

2.2.3 Microwave cavities

The microwave cavities have been previously described [2] and only a short overview is given here. The state-selection cavity and the Ramsey cavity are identical, and the following description applies to both cavities. The cavities and the flight tube above them serve double duty as the vacuum wall. This design suppresses problems with microwave leakage.

Each cylindrical cavity operates at 9.192 631 GHz in the TE_{011} mode and has a 3 cm radius. The cavity height is approximately 2.18 cm. Two quarter-wave chokes suppress the unwanted TM_{111} mode, which is normally degenerate with the TE_{011} mode. The cavity is fed magnetically in the mid-plane of the cavity by four equally spaced circular apertures with a diameter of approximately 0.5 cm. The four apertures couple energy from a resonant mode-filter. The cavity is severely under-coupled. The theoretical unloaded cavity Q for this cavity is 22 000 and, as a result of the small coupling, the loaded Q is nearly equal to the unloaded Q. Atoms enter and leave the cavity through long (8 cm) below-cutoff waveguides of 1 cm diameter, which are centred on the cavity diameter. The cavity was designed with the reduction of distributed-cavity phase shifts in mind and this system is further discussed in Section 4.

2.2.4 Detection system

The detection system consists of two regions: the first (upper) detects $F=4$ atoms and the second (lower) $F=3$ atoms. The two regions are identical with respect to the detection systems for the atomic fluorescence, but they differ in the details of the optical interrogation beams used.

Each fluorescence detection system uses a large (approximately 10 cm diameter) spherical mirror and an optical telescope to image the fluorescence light on to a large-area silicon photodiode. The solid angle for light collection is (1.5 ± 0.15) sr. The detection-system electrical noise is less than the signal from 10 atoms. Noise in the detection process is usually limited by scattered light from the detection beams and the overall noise in the detected normalized signal is equivalent to the signal from 35 atoms. The detection system appears to reach the quantum projection noise limit at around 2500 atoms. The normalization system typically lowers the noise by less than 20 % when the fountain is working well. In the past we have seen much larger reductions when the shot-to-shot atom noise was much worse.

Detection of $F=4$ atoms is accomplished using a σ^+ standing wave tuned near the $|4, m\rangle \rightarrow |5', m\rangle$ cycling transition, where F' is the $^3P_{3/2}$ state. The standing wave is 1 mm high by 20 mm wide $(1/e^2)$ and typically has a saturation parameter of 2.5. Directly below this standing wave is a travelling wave tuned to the same transition, which removes the $F=4$ atoms so that they are not detected in the $F=3$ detection region. Measurements made under operating conditions reveal that more than 99.9 % of the (unpumped) $F=4$ atoms are removed by this beam. Some small fraction of detected $F=4$ atoms, generally less than 3 %, are pumped into the $F=3$ state and detected in the lower detection region.

$F=3$ atoms are detected using the same cycling transition as for the $F=4$ atoms. The atoms are first optically pumped into $F=4$ by a standing wave tuned to the $|3, m\rangle \rightarrow |4', m\rangle$ transition, and the atoms are then detected in a standing wave similar to the one described for the $F=4$ detection system.

Both the $F=3$ and $F=4$ detection systems are above the optical-molasses region. The two detection levels are separated by 9 cm, with the lower one about 15 cm above the molasses region. Graphite in the tube between the trapping and detection regions getters caesium from the high-pressure molasses region. Under typical operating conditions, atoms in the $F=3$ (lower) detection region scatter 10 % fewer photons than $F=4$ atoms as a result of the higher average vertical velocity of atoms in the lower detection level.

2.3 Optical system

The optical system, illustrated in Figure 3, uses two lasers and an optical amplifier. The extended-cavity master laser injects an optical power amplifier which delivers about 300 mW of optical power. The master laser is locked 160 MHz to the red of the $4 \rightarrow 5'$ D_2 saturated absorption transition at 852 nm. The output of the optical power amplifier, at the

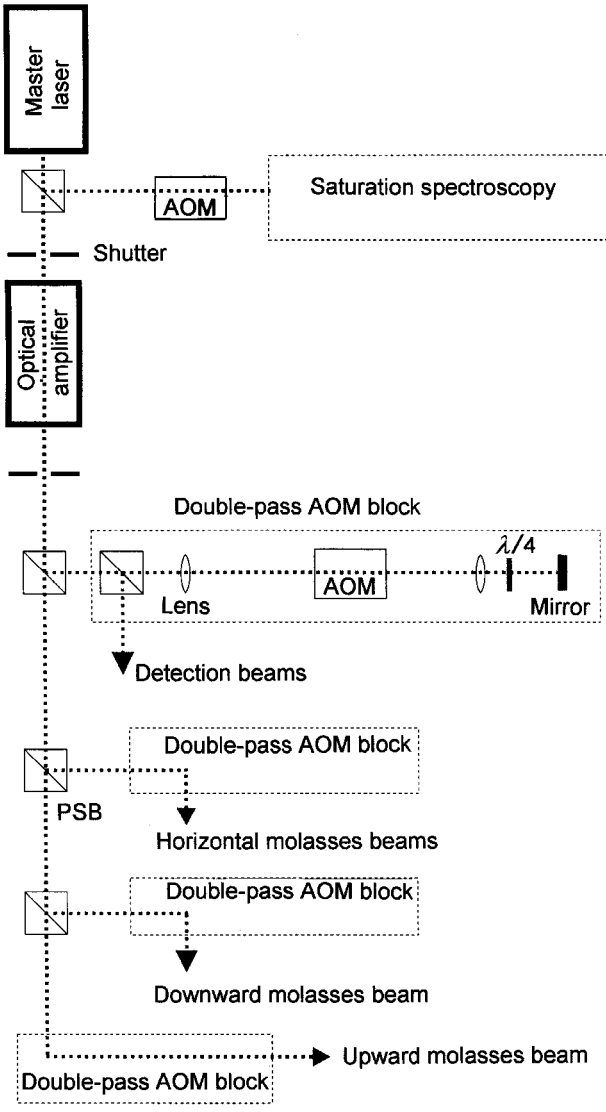

Figure 3. The main laser set-up used in NIST-F1. Not shown are the repumper DBR laser as well as various beam-steering, shaping and polarizing optics.

same wavelength as the master laser, is then frequency shifted to resonance with four double-pass acousto-optic modulators (AOMs) which run at a rf frequency of about 80 MHz. The four AOMs provide the horizontal, up, down and detection beams used in the fountain.

The low power (approximately 5 mW) repump laser is a distributed Bragg reflector type tuned to the $3 \rightarrow 4'$ D_2 transition at 852 nm. This transition is favourable for repumping, with 93 % of the atoms being pumped into $F = 4$ within three photons.

2.4 Control system

The control system for operation of NIST-F1 uses state-machine architecture with states being defined for molasses, launching, post-cooling, state selection, detection, etc. The timing of this state machine is controlled by a digital pattern generator.

The frequency servo part of the control system starts the state machine through one launch cycle, measures the population in the $F = 3$ and $F = 4$ states and calculates the normalized transition probability, $N_4/(N_3 + N_4)$, where N_3 and N_4 are the number of atoms detected in states $F = 3$ and $F = 4$, respectively. The servo operates in either the frequency modulation mode, where the frequency is changed by $\pm 1/4$ of the Ramsey fringe width between successive tosses, or in the phase modulation mode, where the frequency is set to the centre of the Ramsey resonance and the phase is modulated by $\pm 90°$ when the atoms are at apogee. It then changes the frequency (phase) of the microwave generator to the other side of the line (or the other phase) and repeats the process. The difference between the normalized transition probability, $N_4/(N_3+N_4)$, on each side of the line (or each phase) provides the error signal used to steer the centre frequency of the servo. The servo is very similar to that used in NIST-7 [3].

2.5 Microwave synthesizer

The microwave synthesizer, shown in Figure 4, is similar to other NIST synthesizers described in the literature [4, 5]. The short-term stability of the synthesizer is limited by the quartz-crystal oscillators used. These crystals, while high quality, are not quiet enough to support the fountain stability at better than the mid $10^{-13}\tau^{-1/2}$ level. Better quartz crystals have been procured and are being evaluated with the expectation that this will improve the short-term stability for large atom numbers. The specifications of the synthesizer — phase noise, long-term stability, etc. — are discussed in the references and not reiterated here; suffice it to say that the synthesizer is not a significant source of error in these measurements.

3. Frequency corrections applied to NIST-F1

3.1 Overview

In order to realize the definition of the second with uncertainty at the 10^{-15} level, our fountain frequency standard must be corrected for four frequency biases: the spin-exchange shift, second-order Zeeman shift, ac Stark shift (black-body shift) and the gravitational red shift. These biases and their associated uncertainties are discussed in this section.

3.2 Second-order Zeeman shift

In NIST-F1 a small, relatively homogeneous magnetic field, known for historical reasons as the C-field, is applied in the region extending from below the Ramsey microwave cavity to well above the top of the atomic parabolic flight. To excite the desired $\Delta F = 1$, $\Delta m_F = 0$ transition, the microwave magnetic field and the C-field must be parallel. The chosen microwave

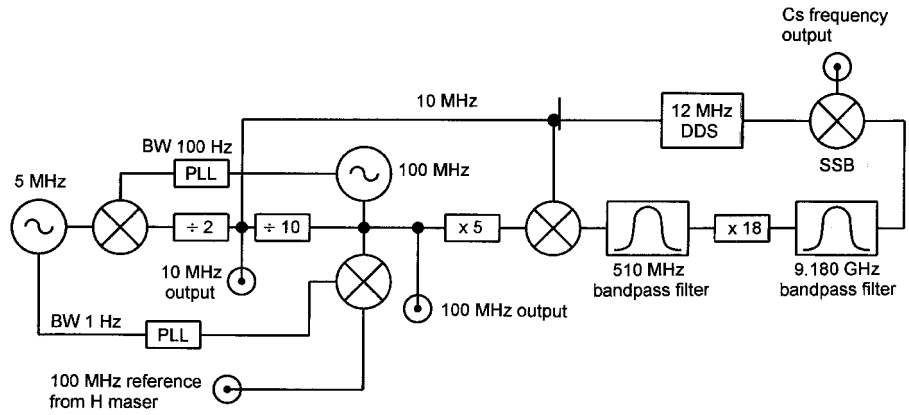

Figure 4. Block diagram of the NIST HR2 microwave synthesizer used with NIST-F1. The output frequency is computer-controlled via the direct digital synthesizer (DDS). BW: bandwidth; PLL: phase-locked loop; SSB: single side-band mixer.

cavity configuration, TE_{011}, has its microwave magnetic field along the z axis. Hence, the C-field is also along the z axis. It typically has a strength of 10^{-7} T (10^{-3} G).

The C-field causes a significant shift of the measured hyperfine splitting of the atoms. To correct for this second-order Zeeman shift, we require knowledge of the magnitude of the magnetic field at all points along the atomic trajectory within and above the Ramsey microwave cavity. This knowledge is embodied by a "map" of the magnetic field at points along the atom trajectory. We first discuss the generation of the magnetic field map in NIST-F1 and then the resulting correction to the frequency along with its associated uncertainty.

3.2.1 Magnetic field mapping

The linear Zeeman effect causes shifts in the measured frequency splittings of the $\Delta F = 0$, $\Delta m_f = 1$ transitions which are directly proportional to the magnetic field. This sensitivity allows the atoms to be used to probe the magnetic field directly. The full map of the magnetic field is constructed from a number of different measurements at discrete locations along the atomic flight region. Each measurement consists of a number of repeated steps which are outlined below.

First the atoms are launched at a velocity so that the apogee of the trajectory is at the desired position for a magnetic field measurement. These atoms are then state selected in the usual fashion (Section 2.1) and the resulting pure $|3, 0\rangle$ sample continues upwards until it reaches apogee. The Ramsey cavity is not excited for these measurements. Instead a low-frequency (roughly 350 Hz) magnetic field transverse to the flight direction is energized. This field excites transitions to other m_F sublevels while leaving the $F = 3$ hyperfine level unchanged. The low-frequency field is turned on shortly (50 ms, or about 1 cm) before apogee and stays on until 50 ms after apogee. The sample of $F = 3$ atoms now falls back through the state-selection cavity which transfers any remaining $|3, 0\rangle$ atoms into the $|4, 0\rangle$ state. The relative signal strengths in the $F = 3$ and $F = 4$ levels of the detection system are then combined to give a transition probability for the atoms out of $m_F = 0$ at the particular frequency and strength of the oscillating magnetic field applied near apogee. The frequency of this low-frequency magnetic field is incremented by a small amount (typically 0.2 Hz) and the process repeated. In this way a resonance curve such as that shown in Figure 5 is generated. The centre of the resonance, which can be identified to within about 0.1 Hz, is a measure of the magnetic field at a particular apogee. The apogee height is now incremented by a small amount, typically between 1 cm and 3 cm, and the resonance curve remeasured. This entire process is repeated for apogee heights from just above the Ramsey cavity to 80 cm above the cavity where the atoms are almost exiting the magnetic shields.

The magnetic field in the Ramsey cavity is measured by state selecting $|3, 1\rangle$ atoms in the state-selection cavity and then exciting the $|3, 1\rangle \rightarrow |4, 1\rangle$

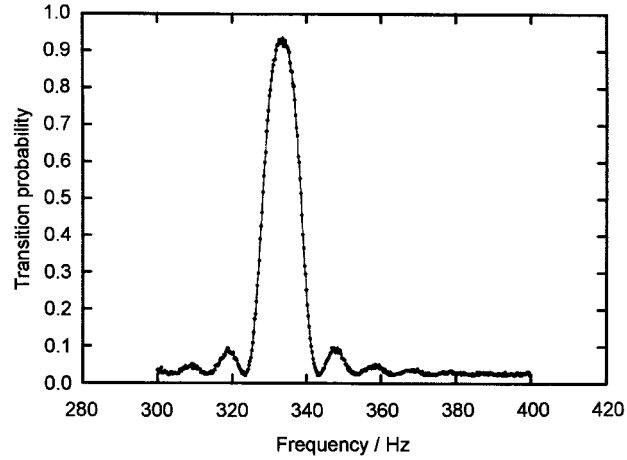

Figure 5. Resonance curve of the $|3, 0\rangle \rightarrow |3, 1\rangle$ transition obtained through the low-frequency resonance technique. The centre of the modified Rabi curve can typically be found with a frequency uncertainty of less than 0.1 Hz.

magnetic-field-sensitive transition with a single π pulse in the Ramsey cavity. The measurement of the field in the Ramsey cavity typically has a 0.2 % relative uncertainty.

The frequency data gathered using the procedures just described are converted into magnetic field units with the aid of published magnetic-field-sensitivity coefficients: 350.98×10^7 Hz/T for the low-frequency $F=3$, $\Delta F=0$, $\Delta m_F=1$ transition and 700.84×10^7 Hz/T for the $\Delta F=1$, $\Delta m_F=0$ transition in the Ramsey cavity [6]. Higher-order terms in the magnetic-field sensitivity are not required at the necessary level of precision.

In this way a magnetic-field map such as that shown in Figure 2 is generated. The uncertainties associated with each datum in Figure 2 include the positional uncertainty at apogee (approximately 0.3 cm) and the uncertainty associated with the centre of the resonance curve (approximately 0.1 %). These uncertainties are smaller than the symbols used in the graph. Systematic biases, including the Millman effect, are estimated to be smaller than the statistical uncertainties just identified.

In principle, the second-order Zeeman shift of the $|3,\ 0\rangle \rightarrow |4,\ 0\rangle$ hyperfine transition can be directly evaluated by calculating the time average of the square of the magnetic field, $\langle B^2(t) \rangle$ over the atomic trajectory. The published coefficients for the second-order Zeeman effect then give a frequency shift of $\delta\nu = (427.45 \times 10^8\ \text{Hz/T}^2) \times \langle B^2(t) \rangle$. The shift is about 400 μHz at the C-field strengths typically used in NIST-F1 [6]. This method of calculating the frequency shift directly from the measured field map is robust and can be expected to have residual fractional frequency errors of the order of 10^{-16} at the field strengths used here. However, as a double check, we extend our investigation of the magnetic field further, as detailed below.

3.2.2 Ramsey fringes, field maps and overlays

The Breit-Rabi formula, shown below, predicts the behaviour of the transition frequency between the ground-state hyperfine levels with the imposition of an external magnetic field. For the hyperfine transitions of immediate interest ($\Delta F=1$, $\Delta m_f=0$) the Breit-Rabi formula can be written (up through second order in the magnetic field)

$$\nu_{(F=3,\,m_f)\Leftrightarrow(F=4,\,m_f)} = \nu_0 + 700.83 \times 10^9\, m_f H_0 + 26.715 \times 10^8\,(16 - m_F^2)\, H_0^2, \quad (1)$$

where ν_0 is the unshifted caesium hyperfine transition frequency (in Hz), m_f is the projection of the angular momentum vector F along the magnetic field H_0 (in T). The last term in (1) is the second-order Zeeman shift. It is somewhat clearer to write the Breit-Rabi formula in terms of the dimensionless parameter x as

$$\nu_{(F=3,\,m_F)\Leftrightarrow(F=4,\,m_F)} = \nu_0\left[1 + \frac{1}{4} m_F\, x + \frac{1}{2}\left(1 - \frac{1}{16} m_F^2\right) x^2 + \cdots\right], \quad (2)$$

where x is Ramsey's dimensionless field parameter [7] ($x = (g_j - g_i)\mu_B H/(h\nu_0)$). The position of the central Ramsey fringe in the $|3,\ 1\rangle \rightarrow |4,\ 1\rangle$ manifold can be predicted with the use of (2) and the magnetic-field map. To simulate what the atoms actually experience, we average over the atom's flight time. This time average heavily weights the region near apogee while giving less weight to the region near the cavity. For example, the 20 % of the spatial distance closest to the cavity only contributes 11 % of the time average. Thus the field inhomogeneity near the cavity is relatively less important at high toss heights.

The magnetic field inhomogeneity shown in Figure 2 causes the frequency of the predicted central fringe to vary strongly with apogee height. Further, it prevents the central Ramsey fringe of the $|3,\ 1\rangle \rightarrow |4,\ 1\rangle$ transition from being centred on the underlying Rabi pedestal. This causes ambiguity as to which fringe is central. The ability to predict the frequency of a particular Ramsey fringe in the $|3,\ 1\rangle \rightarrow |4,\ 1\rangle$ Ramsey manifold therefore serves as a check on the field map previously generated. Figure 6 shows the measured fringe position as a function of apogee height above the Ramsey cavity for several fringes in the $|3,\ 1\rangle \rightarrow |4,\ 1\rangle$ manifold. Also shown in Figure 6 is the predicted location of the central fringe generated from the field map in Figure 2. The difference between the predicted versus measured position is quite small. The statistics of predicted minus measured fringe position have an average of 0.04 ± 0.05 fringes, consistent with zero. Further, because the fringe spacing varies with toss height, the ability to predict the central fringe location through the gyrations shown in Figure 6 provides further evidence that the central fringe has been located. To illustrate this, the field map in Figure 2 was rescaled in order to move the predicted central fringe into agreement with the +1 fringe at the highest toss height. The new "predicted central fringe" now agrees only at the highest toss height and disagrees significantly (calculated fringe minus measured fringe statistics give 0.40 ± 0.06 fringes) at all other toss heights. This "fit" is shown by the grey line in Figure 6. Similar results apply for rescaling to the other fringes.

While not shown in Figure 6, we have traced the predicted central fringe from just above the Ramsey cavity upwards. This technique, as discussed in [8], is based on the fact that at sufficiently low toss heights there are only a few (three) Ramsey fringes and which fringe is central is unambiguous. The central fringe identified using this technique agrees with the fringe predicted using a numerical integration of the field map.

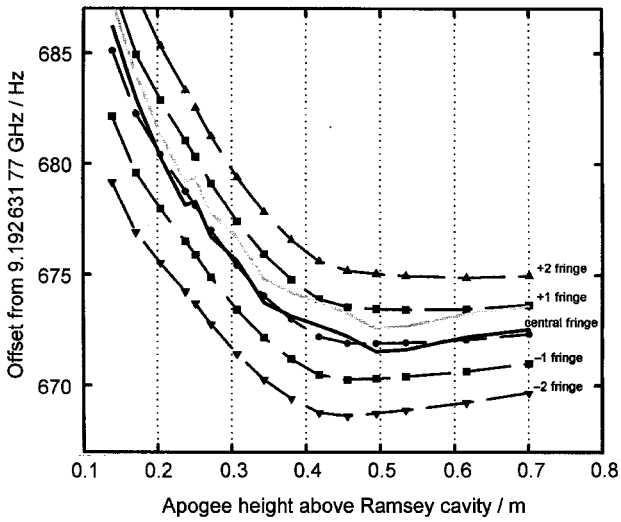

Figure 6. The various symbols connected by dashed lines are the position of Ramsey fringes from the $m = 1$ hyperfine manifold as a function of distance above the Ramsey cavity. The solid black line represents the predicted position of the central fringe from the time integral of the magnetic field in Figure 2 for apogee at various heights above the Ramsey cavity. The agreement between the predicted and measured fringe is quite good, with an average discrepancy of 0.04 ± 0.05 fringes. The grey line results from rescaling the field map in Figure 2 to get agreement with an adjacent fringe. This is not successful, with the predicted fringe missing the measured fringe by an average of 0.4 ± 0.05 fringes. See Section 3.2 for a complete discussion.

In NIST-F1 the numerical integration is considerably less time-consuming than laboriously following the central fringe from an apogee 0.5 cm above the Ramsey cavity to a final apogee at 50 cm or more.

While there is no reason to believe that the central fringe has been misidentified, we assign a (perhaps overly) conservative error of $\delta\nu/\nu = 3 \times 10^{-16}$ in the uncertainty budget shown in Table 2. The fractional uncertainty assigned is equivalent to one full fringe at the field values used in NIST-F1.

Once the central fringe on the $|3, 1\rangle \rightarrow |4, 1\rangle$ manifold has been identified it can be used to predict the frequency offset of the $|3, 0\rangle \rightarrow |4, 0\rangle$ transition due to magnetic field. Let $\Delta\nu_{1,0}$ be the frequency difference between the $|3, 1\rangle \rightarrow |4, 1\rangle$ and $|3, 0\rangle \rightarrow |4, 0\rangle$ transitions. Then, from (2), the frequency correction due to the magnetic field (second-order Zeeman shift) on the $|3, 1\rangle \rightarrow |4, 1\rangle$ transition can be written

$$\Delta\nu_{Q,z} = \frac{8\,\Delta\nu_{1,0}^2}{\nu_0}. \qquad (3)$$

There are several reasons for using this method to correct for the second-order Zeeman shift. First, the use of (3) suppresses a small bias due to underlying pedestal shifts as discussed in [3]. Second, the frequency of the central fringe can be monitored to observe (and if necessary correct for) time fluctuations of the magnetic field. The magnetic field in NIST-F1 has been monitored using this technique and Figure 7 shows typical results. The large bump in the Allan variance plot around 5×10^4 s is consistent with the observed diurnal fluctuation of the Earth's magnetic field. The fractional frequency fluctuation on the $|3, 0\rangle \rightarrow |4, 0\rangle$ transition caused by this feature is less than 10^{-17} and is not currently corrected for.

Table 2. Known frequency biases and their associated Type B uncertainties. This is the actual bias list submitted to the BIPM for the June/July 2001 frequency evaluation. The range of measured spin-exchange biases is shown in parentheses for the spin-exchange bias.*

Physical effect	Bias magnitude /10^{-15}	Type B uncertainty /10^{-15}
Second-order (quadratic) Zeeman	+44.76	0.3
Second-order Doppler	< 0.1	< 0.1
Cavity pulling	< 0.1	< 0.1
Rabi pulling	< 0.1	< 0.1
Cavity phase (distributed)	< 0.1	< 0.1
Fluorescent light shift	< 0.1	< 0.1
Adjacent atomic transitions	< 0.1	< 0.1
Spin exchange	0.0*(0.7 to 7.8)	0.84
Black body	−20.6	0.3
Gravitation	+180.54	0.1
Electronic shifts		
RF spectral purity	0	< 0.1
Integrator offset	0	< 0.1
AM on microwaves	0	< 0.1
Microwave leakage	0	0.2
Total Type B uncertainty		0.99

*See Section 3.5.2 for a discussion of the reporting method for this bias.

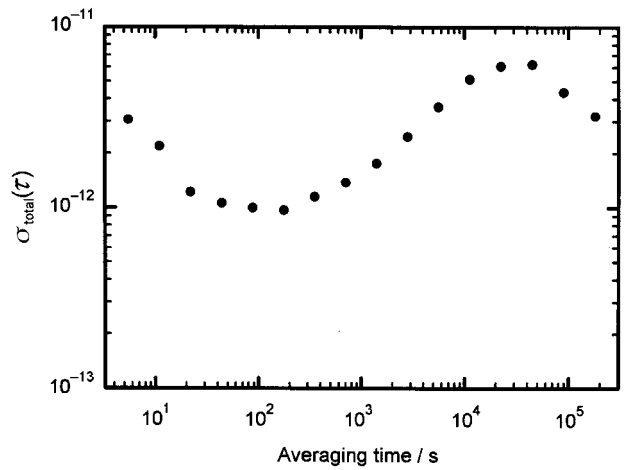

Figure 7. The frequency stability, total deviation, of NIST-F1 when locked to the $|3, 1\rangle \rightarrow |4, 1\rangle$ magnetically sensitive transition. The bump between about 2×10^4 s and 4×10^4 s is a diurnal magnetic field fluctuation. The fluctuation is about 1.6×10^{-12} T and corresponds to a fractional frequency fluctuation on the $|3, 0\rangle \rightarrow |4, 0\rangle$ transition of less than 10^{-17}.

A possible frequency bias results from the magnetic field inhomogeneity and the use of (3). The frequency offset $\Delta\nu_{1,0}$ is, in essence, a measure of the time average of the magnetic field as seen by the atoms, $\langle B(t)\rangle$. This average is then squared with the use of (3) and the squared average, $\langle B(t)\rangle^2$, is used in place of $\langle B^2(t)\rangle$. This type of error is discussed more fully in [3]. In NIST-F1 the fractional frequency bias from this effect is less than 10^{-16} and is not corrected for.

3.3 Black-body radiation frequency shift

The ac Stark shift from the thermal radiation environment of the atoms, first calculated by Itano et al. [9], has recently been measured by two separate groups [10, 11]. The experimental results combined with theoretical results give a frequency shift of

$$\delta\nu = -1.573\,(3) \times 10^{-4}\,\mathrm{Hz} \times \left(\frac{T}{300\,\mathrm{K}}\right)^4 \times \left[1 + 0.014\left(\frac{T}{300\,\mathrm{K}}\right)^2\right], \qquad (4)$$

where T is the temperature (in kelvins) of a perfect black body producing the radiation. The atomic-flight region of NIST-F1 is temperature controlled at about 41 °C. There is a small temperature gradient between the Ramsey cavity and atom apogee of less than 0.5 °C. The flight path of the atoms is well shielded from the outside world and the thermal radiation inside the standard should therefore be representative of the temperature measured by the thermocouples attached to the flight tube. The average wall temperature seen by the atoms is only very slightly shifted by the temperature gradient because the atoms spend the vast majority of the time in the apparatus near apogee where the temperature is essentially constant. With these data in hand the black-body radiation shift can be calculated to be $\delta\nu/\nu = -20.6 \times 10^{-15}$ with an uncertainty of 0.3×10^{-15}, corresponding to a 1 °C uncertainty in the radiation temperature.

The temperature of the radiation field inside the drift tube is somewhat uncertain, but the optically coated window at the top of the drift tube is within 5 °C of the drift-tube temperature and should be relatively "black" in the infrared. Furthermore the solid angle subtended by this window is small ($< 2 \times 10^{-3}$ sr) as "seen" by the atom ball. The radiation field sampled by the atoms should thus be well characterized by the wall temperature. The room is kept dark when the standard is being operated.

3.4 Gravitational frequency shift

The altitude in Boulder, CO, where NIST-F1 is located, is about 1600 m. General relativistic calculations predict a fractional increase in the frequency of a clock when operated above the rotating geoid of 1.09×10^{-16} m^{-1}. In order for NIST-F1 to be compared with and contribute to TAI the clock frequency must be corrected to the geoid. The clock frequency reported therefore has a fractional frequency correction of 180.54×10^{-15} with an uncertainty of 1×10^{-16} [12].

3.5 Spin-exchange frequency shift

The spin-exchange shift in a caesium fountain is extraordinarily large [13, 14] as well as being energy dependent [15] and must be corrected to enable caesium fountains to produce accuracies of better than a few parts in 10^{15}. The spin-exchange coefficient has been measured by several groups [13, 14]. The current understanding that the spin-exchange shift is energy dependent makes the measurements reported in [13, 14] somewhat uncertain. As a result of this uncertainty, we have switched from a method which used the spin-exchange coefficient as reported in [13, 14] along with an absolute density determination (used in the first two formal accuracy evaluations) to a density extrapolation method described next. It is important to note that, within the stated uncertainties of the two methods, the results of the spin-exchange correction using the previously reported spin-exchange coefficient and absolute density determination are in agreement with the results obtained by extrapolation to zero density.

3.5.1 Extrapolation to zero density

The spin-exchange frequency shift for collisions between caesium atoms in the microkelvin temperature range is expected to scale linearly with the atomic density [13-16]. We therefore operate the fountain with otherwise constant parameters (launch height, atom temperature, physics package temperature, C-field, etc.) and vary the number of launched atoms. The zero atom density intercept of an atom number versus frequency plot is then the frequency at which the fountain would operate in the absence of the spin-exchange shift. Figure 8 shows data gathered using this procedure. The horizontal axis of Figure 8 (signal) is a measure of the number of detected atoms. The vertical axis (relative frequency) is the frequency of the reference maser, corrected to the AT1E timescale (an internal NIST timescale). The timescale stability is discussed in some detail in Sections 5, 6 and 7, and it suffices here to say that the stability of the timescale is less than $\sigma_y = 5 \times 10^{-16}$ over the period when these measurements were taken. A weighted linear mean least-squares fit was used to obtain the intercept, slope and uncertainties.

The atom number for the data in Figure 8 was varied by varying either the caesium oven temperature or the molasses time or both. In addition, the atom number is stabilized by use of a number servo which adjusts the microwave power in the state-selection

cavity so that the average returned atom number is constant (to within 1%). In order to ensure that the atomic distribution was invariant under these changes, we measured the initial spatial distribution along the vertical direction as well as the velocity distribution along that direction. The initial spatial distribution along the launch direction is measured in the detection region (above the atom collection region) shortly after launch (about 30 ms). The spatial distribution of the atoms after a complete round trip through the Ramsey interrogation region is also measured. The ratio of the Fourier transforms of these two signals is the Fourier transform of the velocity distribution along the launch direction. Figure 9 shows typical spatial distributions, as well as the velocity distribution derived from the spatial distribution. We also measured the fraction of launched atoms that returned to the detection region, which is a measure of the constancy of the spatial distribution in the plane perpendicular to launch as well as the atomic velocity distribution in that plane. While it is true that both of these could vary together in such a way as to leave the fraction of returned atoms unchanged, the non-linear variation required must be viewed as highly unlikely. We have looked for this kind of variation in another, similar, fountain apparatus at the NIST and have seen no evidence of such an effect. The measured spatial distributions, velocity distributions and returned atom fractions were constant to well under 5%. The atom number was varied over about a factor of 9 for these tests.

3.5.2 Results of the extrapolation to zero density

In NIST-F1 we currently conduct a full density extrapolation during each accuracy evaluation. The result of this density extrapolation is a single point: the predicted frequency at zero-spin exchange bias along with its associated uncertainty. This is what is reported in the table of biases. Clearly, no actual data

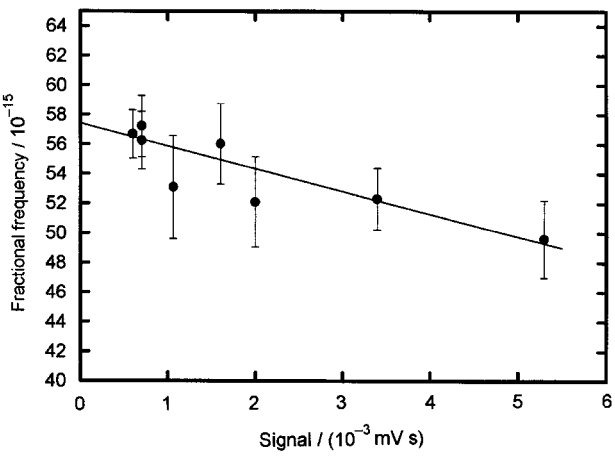

Figure 8. Spin-exchange shift data: the measured fractional frequency of the maser, corrected to the NIST AT1E timescale as a function of detected atom number. The spin-exchange fractional frequency shift varies from about $\mathbf{0.7 \times 10^{-15}}$ to $\mathbf{7.8 \times 10^{-15}}$. The intercept of the curve is the zero-density limit of these measurements. The uncertainty of the intercept is $\mathbf{1.2 \times 10^{-15}}$, of which the combined statistical uncertainty is $\mathbf{0.86 \times 10^{-15}}$ and the systematic uncertainty is $\mathbf{0.84 \times 10^{-15}}$.

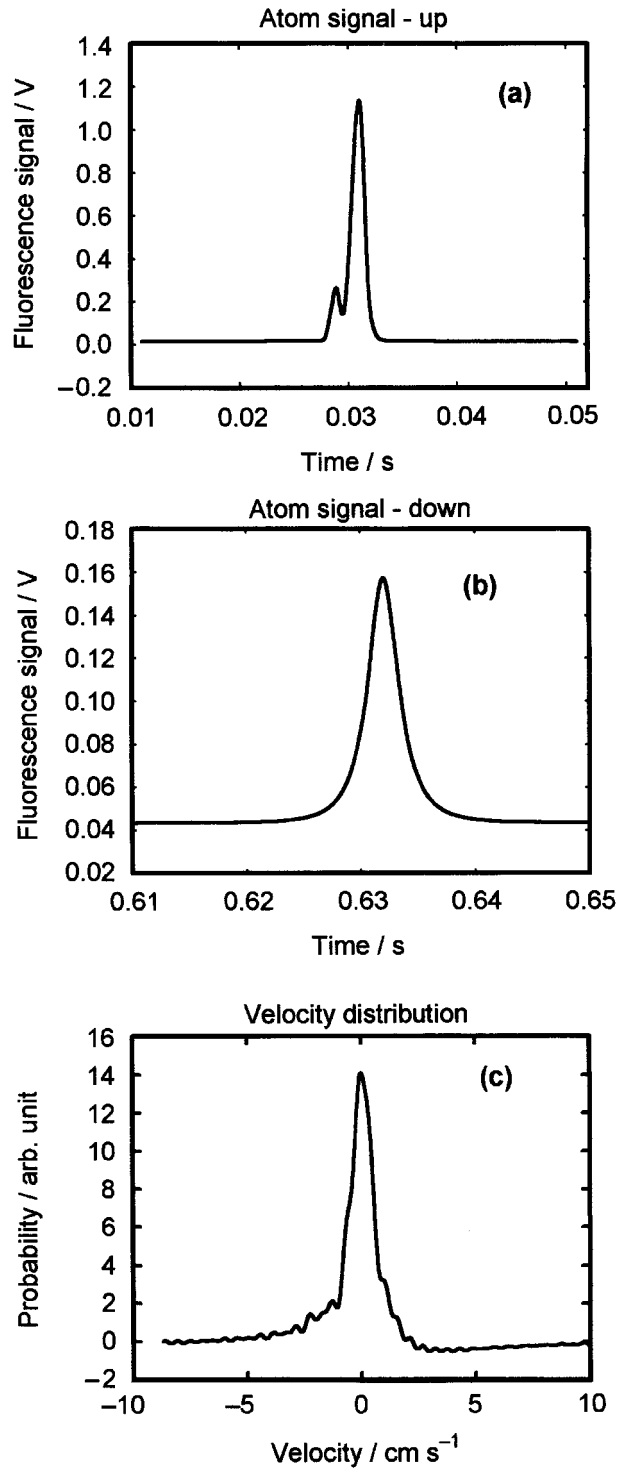

Figure 9. Typical measured data for the density profile from atoms (a) 0.030 s after launch; (b) approximately 1.03 s after launch. These can be combined to yield a velocity distribution, shown in (c).

are gathered at zero density. Because the density is varied over almost an order of magnitude, no unique spin-exchange bias can be reported in Table 2.

Other fountain groups typically measure the slope of the extrapolation curve one or more times and use this historical slope to correct data with a "known" spin-exchange shift. The bias in the latter case would not be zero.

The majority of time gathering the data in Figure 8 is spent on the low atomic density data, which have a spin-exchange shift of less than $\delta f/f = 1 \times 10^{-15}$. The uncertainty of the intercept for the data in Figure 8 is $\delta f/f = 1.2 \times 10^{-15}$, which we regard as a combination of both systematic and statistical uncertainties. We separate these as follows. The intercept uncertainty is regarded as the quadrature sum of the total statistical uncertainty of the points in Figure 8 and an unknown systematic shift; the unknown systematic uncertainty is then the square root of the difference between the square of the uncertainty in the intercept and the square of the statistical uncertainty of the data set. This procedure yields a fractional systematic uncertainty in the spin-exchange shift of 0.9×10^{-15} for the data in Figure 8. The separation of the data in this fashion has the advantage that when the BIPM recombines the Type A and Type B uncertainties the combined standard uncertainty, which is the quadrature sum of the Type A and Type B uncertainties, is correct. The choice is, however, arbitrary and in fact the combined standard uncertainty is the important number.

3.5.3 Other techniques, MOTs vs molasses

We have used only optical molasses in these tests. The use of density extrapolation with a MOT-based source requires, in the light of the present understanding that the spin-exchange coefficient is highly energy dependent [15], a great deal of caution. The diameter of the atomic sample when using a MOT is known to grow with the atom number. The collision energy changes over timescales on the order of the diameter of the initial sample size divided by the "thermal" velocity. With typical MOT sizes of a few millimetres and velocities of about 1 cm/s, this results in the average collision energy being a strong function of the atomic number when using a MOT. As a result of these considerations we use exclusively a molasses which does not suffer from this defect. Note that it should be possible to use a number servo such as that described in Section 3.5.1, with a MOT in which the parameters are held constant and the atom number varied by state-selection power to keep the average collision energy more nearly constant. Even in the case of a "constant" MOT, however, small misalignments of the beams can lead to varying spatial distributions (as a result of imparted angular momentum) without undue effect on the launched number. This leads to a varying average collision energy with time and thus non-constant spin-exchange shift. The effect is likely to be much more pronounced in the case of a MOT than with optical molasses.

4. Frequency biases not corrected in NIST-F1 and their associated magnitudes

4.1 Overview

A large number of biases that are of concern in a traditional thermal-beam caesium frequency standard are considerably reduced in the fountain as a result of the long Ramsey time. Additionally, other possible biases such as electronically caused shifts have been evaluated and an upper limit set on them. These biases, all of which turn out to be either intrinsically small or are kept small through experimental practice, are discussed in this section.

4.2 Doppler shifts of first and second order

The first-order Doppler shift, which manifests itself primarily as distributed-cavity phase shift, is caused by atoms sampling different phases of the microwave field within the Ramsey cavity during their two trips through the cavity. In the best of all possible worlds, this phase shift would be absent as the phase of the microwave fields within the cavity would be constant. The cavities used in NIST-F1 are designed to have a small distributed phase, thereby minimizing the associated frequency shift [2]. The calculated phase extrema for the cavity are of magnitude $0.3\,\mu$rad at the edge of the aperture through which atoms travel. As a (physically unrealizable) worst-case scenario, consider that all the atoms in the atom ball sample a phase of $+0.3\,\mu$rad on the way up through the Ramsey cavity and sample a phase of $-0.3\,\mu$rad on the way down. This would cause a fractional frequency shift of order 2×10^{-17} and is neglected here.

The possibility of a longitudinal phase gradient also exists. In the fountain geometry, time-reversal symmetry causes this effect to be largely cancelled. The residual (uncancelled) portion of the longitudinal phase gradient is proportional (to first order) to the transverse phase along the atomic trajectory divided by the square of the cavity half-height. The phase of the microwave field at the aperture is essentially constant while the phase at the centre of the cavity is more complex, as described in [2]. The number of cavity feeds is also critical in the longitudinal phase in a fountain, because a frequency shift enters as a result of the differential longitudinal phase gradient across the cavity aperture rather than the longitudinal phase gradient. The cavity used in NIST-F1 is almost two orders of magnitude less sensitive to this effect than the cavity described in [17]. The relative frequency shift resulting from residual longitudinal phase gradients in NIST-F1 is less than 10^{-17}.

The second-order Doppler effect is relativistic in nature: moving clocks run slow. An advantage of the fountain over traditional beam standards is that the atoms in a fountain standard move very slowly. The second-order Doppler effect for atoms with a 1 s Ramsey time is of order 3×10^{-17} and is not corrected for, although it could easily be corrected for if necessary.

4.3 Rabi and Ramsey pulling

Rabi pulling, caused by tails of excitation of other $\sigma(\Delta m = 0)$ transitions in the microwave transition spectrum, is almost absent in this standard as a result of state selection. As a worst-case estimate, consider an excitation of only the $m = +1$ σ transition with a height 10 % of the central peak (note that this has never been observed in our clock; typically the $m = +1$ and $m = -1$ Rabi peaks have less than 0.1 % the height of the clock transition Rabi peak). For this worst-case estimate, the fractional frequency shift as a result of Rabi pulling is about 10^{-17}. Using the experimental data measured after state selection, the calculated Rabi pulling is of order 10^{-19}. This shift is not corrected for.

Ramsey pulling is caused by excitation of the $|3, 0\rangle \rightarrow |4, \pm 1\rangle$ π transition $(\Delta m = \pm 1)$, as well as the $|3, \pm 1\rangle \rightarrow |4, 0\rangle$ although the latter transition is not favoured in our state-selected standard. This type of transition is excited if the microwave magnetic field and the C-field are not exactly parallel at all locations where atoms sample the microwave field. This lack of parallelism is guaranteed to occur within our microwave cavity with its half-sine-wave intensity profile. Ramsey pulling is extremely difficult to evaluate completely and we give only a first-order glance at the problem here. Possible coherence effects in the atomic sample have not been evaluated; however, the state-selection process with an optical pulse to remove $F = 4$ atoms should destroy any coherence by projecting the remaining atoms into the $|3, 0\rangle$ state. More complete results will be published elsewhere.

Cutler et al. have given a simplified theory that provides an order of magnitude estimate of the problem [18]. Ramsey pulling is proportional to the difference between the population of the atoms making the $\Delta m = +1$ transition and the population making the $\Delta m = -1$ transition, relative to the population making the clock transition. In NIST-F1 less than 0.1 % of the atoms make a $\Delta m = 1$ transition and the left/right asymmetry is less than 10 %. Given the 10^{-7} T C-field, the approximate order of magnitude of the fractional frequency shift due to the Ramsey pulling terms in [6] is 10^{-17} and is not corrected for.

4.4 Majorana transitions

This fountain is state selected within the magnetic shields. The detection system is essentially insensitive to the m state of the arriving caesium population, being sensitive only to the F state. Therefore, Majorana transitions can cause a frequency shift only if such transitions take place within the magnetic shield structure. Majorana transitions occur if the so-called adiabatic condition is not fulfilled [6]. The adiabatic condition can be written

$$\frac{\beta V}{2\pi l} \sin\theta \ll \Delta\nu_{1,0}, \tag{5}$$

where β describes the field variation, θ is the angle between B (the magnetic field vector) and $\partial B/\partial z$, l is the length scale of the field variation, and V is the velocity of the atoms through the field variation. This condition is satisfied in NIST-F1 by more than three orders of magnitude. The probability is that not a single atom in the caesium cloud makes a Majorana transition within the magnetically shielded region. Given the measured populations in $m = \pm 1$ of less than 10^{-3} (limited by S/N) relative to the $m = 0$ population after state selection and assuming a 100 % left/right asymmetry yields a potential frequency shift of $\delta f/f \approx 2 \times 10^{-16}$ [19]. Additionally, as explained in [19], proper choice of the C-field value suppresses this shift further. The C-field value in NIST-F1 is carefully chosen to satisfy the conditions for suppression of frequency shifts due to Majorana transitions so that the potential $\delta f/f \approx 2 \times 10^{-16}$ is suppressed further by more than an order of magnitude. This possible frequency shift is not corrected for. Zeeman coherences induced by the magnetic field inhomogeneity should not in principle cause a frequency shift. However, as a test we applied a 100 ms pulse of resonant low-frequency excitation to the sample about apogee (as described in Section 3.2) during an otherwise normal measurement of the fountain frequency. No frequency shift was measured with a resolution of $\delta f/f \leq 8 \times 10^{-15}$, even though the applied pulse was sufficient to induce a 25 % coherence and is more than 10^3 times more likely to induce coherence than the magnetic field inhomogeneity shown in Figure 2.

4.5 Spectral impurities and microwave leakage

The microwave spectrum has been measured and from these measurements a worst-case frequency offset of 3×10^{-18} is predicted. As an additional operational test, the fountain is operated in a range of microwave powers above optimal: as much as 120 times optimal power (11 $\pi/2$ pulses rather than $\pi/2$). No microwave-power-dependent frequency shift has been measured at any elevated microwave power level. Assuming that we are well below saturation in any power-dependent effect, any shift should be linear in the microwave field. Although the resolution of the measurement is limited by the measurement times used, this still places a limit on the effects of spectral impurities of about 2×10^{-16} under the assumption of linearity of the shift with field. In particular, we have never observed

a power-dependent frequency shift such as reported in the LPTF fountain, in spite of repeated attempts to do so [20].

The design of NIST-F1 is intrinsically well shielded from microwave leakage. However, the high-power test just described also tests for a possible microwave leakage error. Frequency shifts as a result of microwave leakage in the fountain geometry are also mainly cancelled as a result of the time-reversal symmetry of the atomic flight through the Ramsey interaction.

4.6 DC Stark effect

The entire microwave interaction region and the drift tube above it are constructed from oxygen-free high-purity copper. Further, temperature gradients along this structure are minimized by active control. The structure is allowed an electrical connection to the outside world at only one place. In view of all these precautions, it is unreasonable to expect a potential difference of even a volt over the length of the standard. Assuming a potential as large as 1 V, the associated electric field would be of order 1 V/m, leading to a fractional frequency shift of less than 10^{-19}. As a worst case, we can assume a 1 V patch field inside the cutoff drift tube; this could provide electric field strengths of order 100 V/m, leading to a frequency shift of around 5×10^{-17} (the atoms spend a small fraction of the time in this region). A 1 V patch field in the drift tube at the apogee of the atomic flight would lead to a frequency shift of less than 10^{-16}. No corrections are made for the dc Stark effect.

4.7 Cavity pulling

Cavity pulling effects are quite small in this standard in spite of the high cavity Q, $Q_c = 22\,000$. The cavity is essentially unloaded and the resulting cavity Q is essentially the theoretical cavity Q for the size of cavity and material (oxygen-free copper). For small cavity detunings, the second-order cavity pulling effect can be written

$$\frac{\delta\nu_0}{\nu_0} = \frac{8}{\pi^2}\left(\frac{Q_c}{Q_l}\right)^2 b\tau \cot b\tau \left(\frac{\delta\nu_c}{\nu_0}\right), \qquad (6)$$

where Q_l is the atomic line Q, $b\tau \approx \pi/2$, and $\delta\nu_c$ is the cavity detuning [6]. Assuming that the microwave power is within 1 dB of optimum and that the cavity is kept within one cavity linewidth of the transition, the pulling is only about 8×10^{-17}. In actual operation the cavity is tuned to within much less than one linewidth and the microwave power is closer than 1 dB to optimum. We project a worst-case cavity pulling of 3×10^{-17}. This shift is presently uncorrected for.

First-order cavity pulling, as reported in [21], is much less than $\delta f/f \leq 10^{-17}$ over the range of atomic densities used in NIST-F1.

4.8 Resonant light shift

The laser system for NIST-F1 has three mechanical shutters used to prevent unwanted resonant light from reaching the atoms: one on the DBR repumper section, one between the extended cavity master laser and the optical amplifier, and one on the output of the optical amplifier. Additionally, the AOMs are shut off when the atoms are within the C-field region.

To test the effectiveness of the shuttering, the shutter between the master laser and the optical amplifier was removed. In this case the amplifier delivers approximately 30 dB more optical power than when uninjected. In this configuration, 30 dB more resonant light is produced, yet no resonant light shift greater than 10^{-14} has been observed. We therefore estimate that the resonant light shift when running in the normal configuration (three shutters) should be less than 10^{-16}.

The standard is routinely surveyed with an infrared scanner in order to identify scattered light paths that enter the fountain structure. The room lights are extinguished when the standard is being operated and parts of the standard are draped with an opaque cloth cover.

4.9 Servo biases

The servo system used in NIST-F1 is closely related to the NIST-7 servo system described in [3]. The linewidth splitting required in NIST-7 is, as a consequence of its 72 Hz Ramsey fringe width and 5×10^{-15} accuracy, considerably more challenging than the fringe splitting required for NIST-F1, with its 1 Hz Ramsey fringe and 10^{-15} accuracy.

Two classes of error could be associated with the servo. First is a sloping baseline under the Ramsey fringe. The analysis for this type of error is similar to the analysis required for cavity-pulling effects; for a 10^{-16} frequency bias an effective Q of 2×10^4 located half of the "resonance" linewidth from the central Ramsey fringe would be required. No such effect has been observed. This effect can also be tested by locking the standard on fringes well removed from the central fringe. For example, we have operated the fountain by measuring the left side of the -10th fringe and the right side of the $+10$th fringe. This causes the effective fringe width to be greatly increased, thereby enhancing the effect of a sloping baseline. No frequency offset is observed under these conditions with sensitivity to a 5×10^{-17} frequency shift on the central fringe. Additionally, phase modulation, which we often use on our fountain, is much less sensitive to this effect than frequency modulation.

The second class of potential servo offsets comes from a possible synthesizer frequency offset. We have compared our frequency synthesizer with a newer design, discussed in [5]. The comparison showed no

frequency offset at the 10^{-16} level, even though the synthesizers operate using completely different topologies and components.

4.10 Background gas collisions

Measured pressure in NIST-F1, along with known conductivities, pressure shift coefficients and residual gas analysis, leads to the estimate of fractional frequency shift from this effect of less than 10^{-16} [22]. Given the uncertainty of this estimate, however, we assign a somewhat larger systematic uncertainty of $\delta\nu/\nu = 1 \times 10^{-16}$.

5. Frequency stability

The first step in the formal evaluation process is to evaluate certain biases. The evaluation of biases is easiest if some "leveraged" test can be performed [3]. The ability to cause the bias to be greatly exaggerated for evaluation allows a greater number of biases to be evaluated in any given interval. In any case, the greater the frequency stability the more easily evaluated (in general) the biases. Figure 10 shows the stability of this standard in its present configuration. The 1 s (inferred) stability of $5.8 \times 10^{-13}\tau^{-1/2}$ is consistent with the stability limit imposed by the combination of the quartz crystals in our synthesizer and the duty cycle of the fountain (Dick effect) [23]. This stability requires about 4 days of continuous operation to reach a fractional frequency uncertainty of 10^{-15}.

The quartz oscillators used in the synthesizer exhibit a flicker floor of $\sigma_y \approx 3 \times 10^{-13}$, resulting in a stability floor at 2.1 s of 4×10^{-13}. We routinely reach a noise limit of about $\sigma_y(\tau \approx 2.5\ \mathrm{s}) = 4 \times 10^{-13}$ at the cycle time. This noise limit does not seem to change even when the atom number is reduced by 50 %.

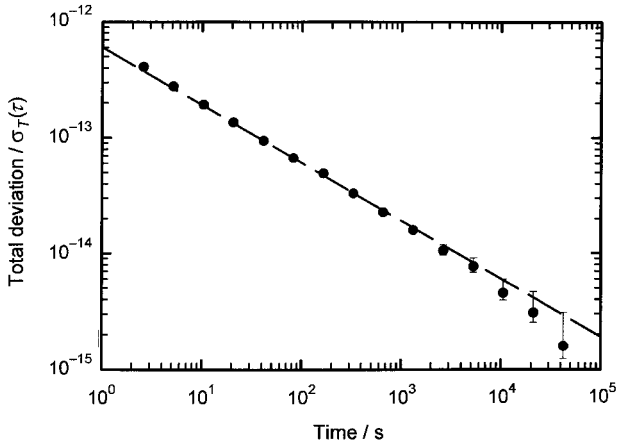

Figure 10. The frequency stability, total deviation, of NIST-F1 when operated as a primary frequency standard. The stability at times longer than shown in the figure is determined by comparison with the NIST AT1E timescale which has a stability of $\sigma_y(\tau) \leq 5 \times 10^{-16}$ for τ between less than 1 day and more than 30 days.

Making a virtue of necessity, this allows us to operate at a very low spin-exchange shift without sacrificing stability. However, better quartz oscillators have been procured and are being incorporated into the fountain.

The great advantage of having an ensemble of hydrogen masers available at the NIST (see Section 7) is the ability to combine discrete data sets in order to extend the effective run time. The NIST post-processed time scale, AT1E, exhibits stability of better than $\delta f/f \leq 5 \times 10^{-16}$ from less than 1 day to more than 30 days [24]. This stability allows individual frequency measurements from NIST-F1 to be combined with a resulting uncertainty less than the uncertainty from any given measurement in the set. Assuming a perfect reference, the resulting uncertainty of N identical measurements would be $1/\sqrt{N}$ times the uncertainty in a single measurement. As an example, the spin-exchange extrapolation evaluation of NIST-F1, made in August and September 2000, had 21 separate frequency measurements made over the course of 29 days, with a total run time of 16.6 days. The smallest uncertainty for an individual run was $\delta f/f = 2.3 \times 10^{-15}$. The combined statistical uncertainty of the entire data set (assuming a perfect reference) was $\delta f/f = 7.7 \times 10^{-16}$. The additional uncertainty caused by the instability of AT1E was $\delta f/f = 2 \times 10^{-16}$ [25], a negligible contribution to the total statistical uncertainty. With the availability of a very stable frequency reference such as AT1E, the statistical uncertainty of an evaluation can be significantly smaller than the last point of an Allan variance plot from an individual run.

6. Formal evaluation procedures

NIST-F1 is not run as a clock. A formal evaluation for the BIPM is conducted by measuring the average frequency of a reference hydrogen maser with NIST-F1 over a specified period. The maser is one of the NIST clocks reported to the BIPM and therefore provides a means by which the maser frequency can be related to the frequency of TAI. From this information the frequency of the fountain relative to TAI is determined.

The biases are evaluated by carrying out a number of tests. A frequency measurement using a range of high microwave powers (typically from $3\pi/2$ to $15\pi/2$) is performed and the results compared with a normal power run (nominally $\pi/2$). This test is also repeated on the state-selection cavity where up to 7π instead of the normal π pulse is applied. These tests check for a number of possible systematics, as discussed in the previous sections. Tests for optical shifts from light leakage are also performed: a shutter is removed as described in Section 4.8 and the resulting frequency shift measured. The fountain is also surveyed with an infrared imaging system to detect possible stray light paths into the apparatus. Next, the magnetic field is mapped in order to evaluate the correction to be

applied for the quadratic Zeeman shift. The atomic spatial distribution and temperature as a function of atom number are checked for constancy over a range of molasses load time and Cs oven temperatures. This allows extrapolation of the spin-exchange frequency shift as a function of the number of Cs atoms.

The standard is then run to measure the frequency of the reference maser, which is in turn referenced to the AT1E timescale. This process may take several days, as the fountain is operated long enough to bring the statistical uncertainty down to an acceptable level. The atomic density is then varied and the above measurement repeated in order to construct a plot of atom number versus measured frequency. These data are eventually used to determine the frequency of the maser corrected for the spin-exchange shift. Because of the high stability of the timescale $(\sigma_y(1\text{ day} < \tau < 30\text{ days}) \leq 5 \times 10^{-16})$ the fountain does not need to run continuously during this period. The most important parameter is total accumulated run time. A significant amount of dead time can be tolerated without seriously affecting the overall uncertainty of the evaluation [25, 26]. During the course of the frequency measurement, the environmental parameters of all five reference masers (temperature, relative humidity, vertical magnetic field, barometric pressure and power line voltage) are monitored along with their frequency stability (see Section 7).

If any significant change is seen in operational parameters of the fountain over the course of the measurement, some or all of the bias evaluations are repeated. Most of the tests described in Sections 3 and 4 are repeated routinely and can be considered part of a formal evaluation. The final step is to apply the measured bias corrections as listed in Table 2 and report the corrected frequency to the BIPM. The biases and uncertainties in Table 2 are those reported to the BIPM for the fifth formal evaluation of NIST-F1. The statistical uncertainty for this evaluation was 8.9×10^{-16}, giving a combined uncertainty of 1.3×10^{-15}. It should be realized that Table 2 lists all *known* biases for NIST-F1; unknown biases cannot, by definition, be corrected.

7. Performance assessment

7.1 Introduction

In addition to the careful evaluation of all biases, as described above, it is also useful to assess the long-term (run-to-run) stability of the standard. This can be accomplished if a sufficiently stable frequency reference is available. A clear indication is given that something is not under control if the long-term frequency stability of a standard is not consistent with the combined uncertainty of the standard and the stability of the reference. Frequency comparisons against other primary frequency standards should also be made.

7.2 Internal comparisons

The NIST is fortunate to have an ensemble of five active, cavity-tuned hydrogen masers and four high-performance commercial caesium-beam tube standards. This ensemble [27, 28], is used to produce a post-processed scale, AT1E [24], that exhibits a frequency stability of 4×10^{-16} at 10 days and which is better than 1×10^{-15} over the range 0.2 to 100 days. The frequency drift rate of AT1E is of the order of -3×10^{-15} per year.

A frequency reference with this high stability allows high-precision comparisons to be made over time. Only one maser at a time is used as the actual measurement reference for the fountain, but the ensemble provides a secondary reference. The fountain frequency for each evaluation can be referenced to AT1E to evaluate its stability, and AT1E also provides information on the stability of the reference maser throughout the course of an evaluation. Therefore, during fountain testing it is easily determined if an anomalous frequency offset is caused by a problem with the fountain or with the reference maser. It is also possible to compare fountain measurements separated by a period of two to three months at the 2×10^{-15} level. Figure 11 shows the fractional frequency offset of AT1E relative to six primary frequency standards, including NIST-F1, over a period of about 1600 days. The horizontal axis is in terms of the Modified Julian Day (MJD). Representative uncertainties are shown for each standard. Only one set of error bars is shown for a standard if the uncertainty is substantially unchanged over the period shown in the figure. For internal frequency measurements with NIST-F1 or NIST-7 (an optically pumped thermal-beam caesium standard), the noise contribution of the NIST measurement system is negligible (at or below 1×10^{-16} for $\tau > 1$ d).

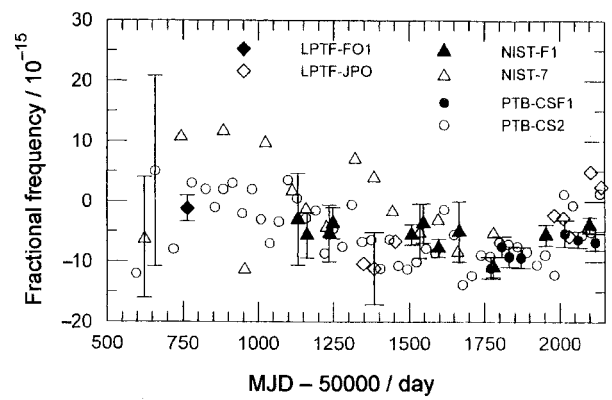

Figure 11. Frequency of AT1E compared with six primary frequency standards over a period of about 1600 days. AT1E is a post-processed maser-dominated ensemble used as a frequency reference.

Eleven NIST-F1 measurements, shown in Figure 11, have been made over a period of about

1000 days and all are statistically consistent within their individual uncertainties and the stability of AT1E. Two different reference masers were used for these measurements. Note that the fountain uncertainties have generally decreased with time as the performance of the fountain has improved. The fifth measurement, made over a period of 4.1 days centred on MJD 51508 (26 November 1999), was the first formal evaluation of NIST-F1, and the result was reported to the BIPM. The seventh, ninth, tenth and eleventh data points were also formal evaluations that were reported to the BIPM. All other data points were informal evaluations and, in most of these cases, the runs were short. This resulted in relatively large statistical uncertainties.

The availability of a stable reference increases confidence in the stated uncertainties when the measurements made over a period of time are consistent. This self-consistency is a necessary, but not sufficient, condition for having correctly determined the overall uncertainty of the standard. Ten NIST-7 evaluations were performed during the period of operation of the fountain and these evaluations are also consistent with the fountain measurements within the stated uncertainties of NIST-7.

7.3 External comparisons

Figure 11 also shows data for the Physikalisch-Technische Bundesanstalt (PTB) thermal-beam caesium standard CS2 and the new PTB fountain standard CSF1 [8]. Two LPTF standards, LPTF-FO1 (a caesium fountain) and LPTF-JPO (an optically pumped thermal-beam standard) are also shown. The data for these standards were taken from *Circular T* and the *Annual Report of the Time Section* of the BIPM, or were obtained through direct communications with the staff at the PTB. Long-distance comparisons are all degraded to some extent by instabilities in the time-transfer technique. Most of the data are for 30-day evaluation intervals using common-view GPS, which gives a time transfer uncertainty of about 1×10^{-15} at 30 days. However, the seven comparisons with PTB-CSF1 were for 15-day (or in the last case 20-day) intervals and were made using two-way time transfer. The two-way link with the PTB has demonstrated greater stability than common-view GPS and gives a time transfer uncertainty of 0.6×10^{-15} for a 15-day interval [29]. The total uncertainties represented by the error bars for the PTB and LPTF standards include the time-transfer uncertainty. The comparisons with the remote standards could also have been made using TAI as the reference.

As may be seen in Figure 11, the NIST-F1 measurements are in agreement within stated uncertainties with other standards that were operated at the same, or nearly the same, time. The seven PTB-CSF1 measurements are particularly important, as they were made over the same period as the last three NIST-F1 measurements and the agreement is good. The approximately 400-day period between the last reported LPTF fountain measurement and the earliest NIST-F1 evaluations makes it difficult to compare these two standards. However, they also appear to be consistent if the quite good PTB-CS2 data are used as a transfer standard.

References

1. Clairon A., Ghezali S., Santarelli G., Laurent Ph., Lea S., Bahoura M., Simon E., Weyers S., Szymaniec K., *Proc. 5th Symp. Frequency Standards and Metrology,* Singapore, World Scientific, 1996, 49-59.
2. Jefferts S. R., Drullinger R. E., De Marchi A., *Proc. IEEE International Frequency Control Symp.,* 1998, 6-8.
3. Shirley J. H., Lee W. D., Drullinger R. E., *Metrologia,* 2001, **38**, 427-458.
4. Nava J. H., Walls F. L., Shirley J. H., Lee W. D., Armburo M. C., *Proc. IEEE International Frequency Control Symp.,* 1996, 973-979.
5. SenGupta A., Popovic D., Walls F. L., *Proc. Joint Meeting European Frequency and Time Forum and IEEE International Frequency Control Symp.*, 1999, 615-619.
6. Vanier J., Audoin C., *The Quantum Physics of Atomic Frequency Standards,* Bristol/Philadelphia, Adam Hilger, 1989, 836.
7. Ramsey N. F., *Molecular Beams,* Oxford, Clarendon Press, 1956, 80.
8. Weyers S., Hübner U., Schröder R., Tamm Chr., Bauch A., *Metrologia*, 2001, **38**, 343-352.
9. Itano W. D., Lewis L., Wineland D., *Phys. Rev. A*, 1982, **45**, 1233-1235.
10. Bauch A., Schröder R., *Phys. Rev. Lett.*, 1997, **78**, 622-625.
11. Simon E., Laurent P., Clairon A., *Phys. Rev. A*, 1998, **57**, 436-439.
12. Weiss M. A., Ashby N., *Metrologia*, 2000, **37**, 715-717.
13. Gibble K., Chu S., *Phys. Rev. Lett.*, 1993, **70**, 1771-1774.
14. Ghezali S., Laurent Ph., Lea S. N., Clairon A., *Europhys. Lett.*, 1996, **36**, 25-30.
15. Leo P. J., Julienne P. S., Mies F. H., Williams C. J., *Phys. Rev. Lett.*, 2001, **86**, 3743-3746.
16. Gibble K., Chang S., Legere R., *Phys. Rev. Lett.*, 1995, **75**, 2666-2669.
17. Laurent Ph., Lemonde P., Abgrall M., Santarelli G., Pereira Dos Santos F., Clairon A., Petit P., Aubourg M., *Proc. Joint Meeting European Frequency and Time Forum and IEEE International Frequency Control Symp.*, 1999, 152-155.
18. Cutler L. S., Flory C. A., Giffard R. P., De Marchi A., *J. Appl. Phys.*, 1991, **69**, 2780-2792.
19. Bauch A., Schröder R., *Ann. Physik,* 1993, **2**, 421-449.
20. Abgrall M. et al., *Proc. 14th EFTF*, 2000, 52.
21. Fertig C., Gibble K., *Phys. Rev. Lett.*, 2000, **85**, 1622.
22. Arditi M., Carver T. R., *Phys. Rev.*, 1961, **124**, 800.
23. See for example the collection of articles in *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, 1998, **45**, 876-905, and references therein.
24. Parker T. E., *Proc. Joint Meeting European Frequency and Time Forum and IEEE International Frequency Control Symp.*, 1999, 173-176.

25. Parker T. E., Howe D. A., Weiss M., *Proc. IEEE International Frequency Control Symp.*, 1998, 265-272.
26. Douglas R. J., Boulanger J. S. , *Proc. 11th European Frequency and Time Forum*, 1997, 345-349.
27. Parker T. E., Levine J., *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, 1997, **44**, 1239-1244.
28. Parker T. E., *IEEE Trans. Ultrason. Ferroelect. Freq. Contr.*, 1999, **46**, 745-751.
29. Parker T. E. et al., *Proc. 15th European Frequency and Time Forum*, 2001, pp. 57-61.

Received on 16 March 2000 and in final form on 14 March 2002.