ARTICLES
https://doi.org/10.1038/s41567-018-0156-2

nature physics

A null test of general relativity based on a long-term comparison of atomic transition frequencies

Neil Ashby[1,2], Thomas E. Parker[1] and Bijunath R. Patla[1*]

The local position invariance principle of general relativity stipulates that non-gravitational experiments should give outcomes that are independent of the position and orientation of the reference frames in which they have been performed. Here, we study the change in the rates of clocks on Earth with the spatial change of the solar potential, constraining the variation of a non-gravitational interaction—the hyperfine splitting in hydrogen and caesium atoms—to $\beta = (2.2 \pm 2.5) \times 10^{-7}$, a factor of two improvement over previous estimates. Our result is based on the comparison between the long-term fractional frequency variation of four hydrogen masers that are part of an ensemble of clocks comprising the National Institute of Standards and Technology, Boulder, and the fractional frequencies of primary frequency standards operated by leading metrology laboratories in the United States, France, Germany, Italy and the United Kingdom over a period of more than 14 years. Using our results together with the previous best estimates of β, we impose strict limits on the variation of fundamental constants, resulting in a test of general relativity with an unprecedented level of precision.

A basic tenet of general relativity is that clock rates and measuring rod lengths are dependent on the amount of energy and momentum in the neighbourhood. Alternatives to general relativity go even further, for example, allowing clock rates to depend on the internal structure of the atoms with which the clocks are constructed, or predicting that the results of similar experiments will differ if performed at remotely located places or times. It is the rates of clocks on Earth that we study here to find any correlation with the spatial variation of the solar potential. Atomic clocks used in this study are so stable that they could be in error by no more than a thousand seconds over the life of our Universe, ~14 billion years, and today, clocks based on optical transitions in ytterbium and strontium atoms perform 1,000-fold better. The analysis presented here provides a recipe for combining the best clock comparisons involving different atomic species for testing one of the main postulates of general relativity, whose future as the fundamental theory of gravity will rest in part on the frequency comparisons using the next generation of clocks.

Several far-reaching principles are embedded in Einstein's general theory of relativity[1,2]. The general consensus is that any metric theory such as general relativity satisfies the Einstein equivalence principle (EEP) that encapsulates three main principles[3]. Local Lorentz invariance (LLI) states that the laws of physics must be independent of the velocity of the reference frame in which the laws are expressed; in other words, the laws of physics must be form-invariant with respect to transformations between relatively moving reference frames. Local position invariance (LPI) requires that the outcome of any non-gravitational experiment be independent of the position and orientation of the reference frame in which the experiment is performed. The weak equivalence principle (WEP) requires that in a gravitational field, all objects—regardless of their internal composition—fall with the same acceleration. LPI is the topic of the present study; the remainder of this paper assumes that both WEP and LLI are valid.

In our study, the hyperfine splitting interactions in hydrogen and caesium atoms arising from magnetic interactions between nuclear and electron magnetic moments are the non-gravitational interactions of interest. We look for variations in atomic transition frequencies arising from such interactions as the Earth orbits the Sun, thereby changing the gravitational potential in which the transitions occur. H and Cs have different internal atomic structures in terms of the neutron to proton ratio (N/Z), and in the electromagnetic contribution to the binding energy ($\propto Z^2$) (ref. [4]). N and Z for H are 0 and 1, whereas for Cs they are 78 and 55. The very dissimilar internal structures of H and Cs atoms should help amplify any LPI-violating signal in a differential measurement.

According to general relativity, if two clocks of different internal structures move together through a gravitational potential, their frequency ratio must be constant, otherwise their frequency shifts relative to a reference at a different gravitational potential would not be unique (see Fig. 1).

Such a comparison does not involve direct time transfer between space-borne clocks and clocks on the ground, nor does it require the clocks to be accurate in frequency. For such tests, the longer-term stability (stability for an orbital period or longer) of the clocks is relevant, and it is clear that the same control of systematic effects that yields high accuracy also leads to high stability. To test relativity to higher precision, accurate measurements of time and frequency—with regard to clock comparisons—are critical. In the past decades, technological advancements in precision metrology have made possible time and frequency measurement with higher precision and better stability[5].

A change in the gravitational potential at the location of a clock, according to various alternatives to general relativity, causes a fractional frequency shift in the clock

$$\Delta f / f = (1+\beta)\Delta\Phi / c^2 \qquad (1)$$

where $\Delta\Phi$ is the change in gravitational potential, Δf is the change in frequency f of the clock and c is the speed of light in vacuum. The parameter β measures the degree of violation of LPI; in general relativity, $\beta = 0$. The eccentricity of Earth's orbit ($e = 0.0167$) provides

[1]National Institute of Standards and Technology, Boulder, CO, USA. [2]Department of Physics, University of Colorado, Boulder, CO, USA.
*e-mail: bijunath.patla@nist.gov

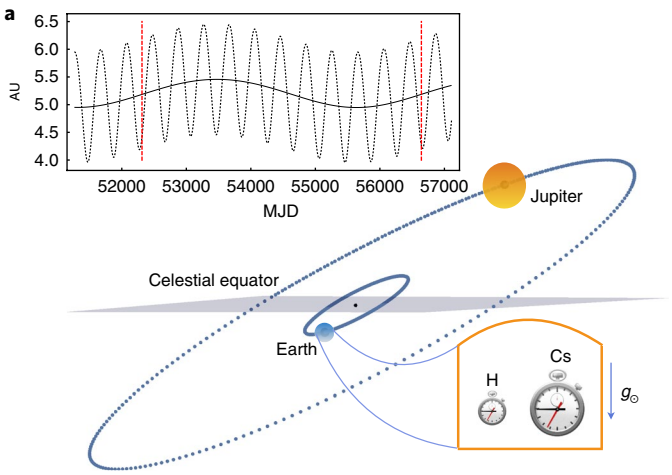

Fig. 1 | Position of Earth and Jupiter on modified Julian day (MJD) 56263 (2 December 2012). Earth may be conceived as an Einstein elevator (neglecting tidal forces). The Cs fountains and the H masers are all considered to be 'local', with spatial separations that are relatively small, in the inertial frame centred on the Earth, as it falls freely in the Solar System's gravitational field with an acceleration $g_\odot \sim 0.006\,\mathrm{m\,s^{-2}}$. This is in compliance with principles of general relativity (or any metric theory) that require LPI to be true only 'locally'. The top panel shows the distance (in AU) of the barycentre of Jupiter from the Sun's centre and the distance between the barycentres of Jupiter and Earth are represented by solid and dotted curves. The time period between the dashed vertical lines is ~11.86 years, the orbital period of Jupiter. 1 AU is roughly the mean distance between the Earth and the Sun, which now has a fixed assigned value (see Supplementary Table 1). Jupiter's radius is ~11 times that of Earth's radius and the orbital radius is ~5.2 times larger than Earth's orbital radius. Planets orbiting the Sun are not drawn to scale.

sufficient variation of the distance separating the Earth and the Sun to assess a possible correlation between the annual variation of gravitational potential and the corresponding frequency offset introduced in the clocks. The size of the Earth is extremely small compared to the variation of the Earth–Sun distance, so the gravitational redshifts arising from the solar potential differences between the two clocks positioned at different locations on the Earth are very nearly the same, and in general relativity are cancelled by relativistic effects arising from free fall. The non-gravitational contribution to the difference in the fractional frequency shifts of two different clock types, H and Cs, is

$$\Delta f/f|_{\mathrm{H}} - \Delta f/f|_{\mathrm{Cs}} \equiv \Delta f/f|_{\mathrm{H-Cs}} = \beta \Delta\Phi/c^2 \qquad (2)$$

where $\beta = (\beta_{\mathrm{H}} - \beta_{\mathrm{Cs}})$. In a null test of a metric theory of gravity such as general relativity, a measurement would put an upper limit on the absolute value of β. While the present work builds on and extends the work of Ashby et al.[6], similar experimental tests of LPI have been a topic of interest for a very long time (see Methods).

The National Institute of Standards and Technology (NIST), in Boulder, Colorado, hosts five H masers and four commercial Cs standards as the basis of the timescale that provides—along with the US Naval Observatory—the official civil time UTC (NIST) for the United States. By international convention, the exact frequency 9,192,631,770 $\mathrm{s^{-1}}$, corresponding to the hyperfine splitting in the ground state, $|F=3; m_F=0\rangle \leftrightarrow |F=4; m_F=0\rangle$, of Cs^{133} atoms held at a temperature of 0 K, provides the international system (SI) definition of the second. The definition is realized at major national laboratories through primary Cs-fountain frequency standards such as NIST-F1 or NIST-F2, which are run intermittently and are used to improve the long-term stability of the NIST timescale, and to help calibrate International Atomic Time (TAI).

For this study, we chose eight such Cs-fountain primary frequency standards from all over the world: IEN-CsF1 from Istituto Nazionale di Ricerca Metrologica, Torino, Italy[7]; NIST-F1 from NIST, Boulder, USA[8]; PTB-CsF1 and -CsF2 from Physikalisch-Technische Bundesanstalt, Braunschweig, Germany[9,10]; NPL-CsF1 and -CsF2 from the National Physical Laboratory, Teddington, UK[11] and SYRTE-CsFO1 and -CsFO2 from Systèmes de Référence Temps-Espace, Paris, France[12]. The fractional frequency shifts of Cs primary frequency standards, referenced to the geoid, are reported to the Bureau International des Poids et Mesures (BIPM), Sèvres, France, ordinarily after each evaluation and are available from the BIPM 'Circular T'[13].

The four NIST H masers that are used in this study are labelled S2 to S5. The masers are housed within environmentally controlled chambers and monitored for fluctuations in pressure, temperature, magnetic field and humidity. The frequency shifts introduced by the environmental variables are computed on the basis of measured frequency sensitivities corresponding to each variable, for each maser. In general, the corrections for environmentally caused frequency shifts are of the order of 10^{-16} to 10^{-13}. For these masers, the temperature corrections were the most consequential and during some epochs were as high as 10^{-13}. The differences in the frequency shifts of fountains versus a typical maser with time, after correcting for changes in environmental variables, are plotted in Fig. 2 and residuals after removing the maser drifts are given in Fig. 3.

The spatial variation of the gravitational potential is computed using DE430 planetary and lunar ephemerides[14]. We calculate the gravitational potential of Jupiter by computing the distance between Jupiter's barycentre and the Earth–Moon barycentre. Jupiter's mass is about a thousand times smaller than that of the Sun but twice that of the rest of the planets in the Solar System. The precision of DE430 ephemerides, which is sub-kilometre within the inner Solar System, allows us to realistically account for the effect of Jupiter's gravity.

For each H maser, the amplitude of the LPI parameter β is computed by using the residuals from the polynomial fits and the combined potential variation due to the Sun and Jupiter for the epoch corresponding to the fountain evaluation. We look for correlation in the residuals with a fixed phase and period corresponding to the variation in the total potential, using equation (2). The uncertainty is obtained by performing a standard least-squares fit (all Cs fountains are assigned equal weights) of the data. The results for the amplitude and uncertainty for all four masers are combined to obtain the final result (see Methods for more details on data analysis)

$$\beta|_{\mathrm{H-Cs}} = (2.24 \pm 2.48) \times 10^{-7} \qquad (3)$$

This study improves the uncertainty in β by more than a factor of five compared to our previous study in 2007 using H and Cs, and imposes a stricter constraint on the uncertainty in β reported for any two pairs of atoms[6,15,16]. The inclusion of Jupiter's gravitational potential had an effect only on the third significant digit with a contribution of the order of a percent. For the entire data set, the uncertainty in the estimation of the gravitational potential variation using the planetary ephemerides is three orders of magnitude smaller than the combined stated uncertainty due to frequency transfer. Therefore, in our calculations leading to equation (3), we have neglected the uncertainty in the estimation of the gravitational potential.

If LPI were violated, with a very small but finite value for β, it would imply that non-gravitational interactions would have also varied with time—resulting in the variation of certain fundamental constants. It is

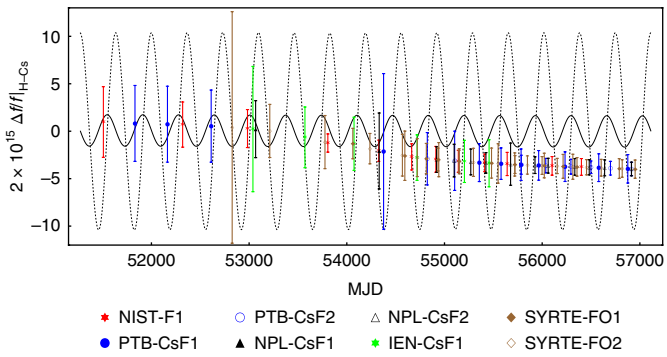

Fig. 2 | Frequency shifts of H and Cs for maser S3 for MJD 51508 (11 November 1999) to MJD 56959 (29 October 2014). The solid curve is the change in gravitational potential, $\Delta\Phi/c^2$, due to the Sun and Jupiter as a function of time and the dotted curve is the total rate of change of gravitational potential, $\Delta\Phi/\mathrm{d}t/c^2[\mathrm{yr}^{-1}]$. Gravitational potential variations are scaled by a factor of 10^{10}, instead of the factor 2×10^{15} (see *y* axis) used for scaling the frequency shifts. The data are the differences in fractional frequency shifts of H and Cs after accounting for environmental corrections for maser S3. Each data point is the average of 10 to 40 days corresponding to the time when the Cs fountains were evaluated. The time stamp of a data point is the midpoint of the evaluation period. The value of the gravitational potential assigned to this midpoint is the average over the fountain evaluation period. The solar potential is evaluated using the distance between the Sun's centre and the Earth–Moon barycentre. Jupiter's potential is calculated using the distance between the Earth–Moon barycentre and the barycentre of Jupiter's system. The uncertainty is the square root of the sum of the uncertainties added in quadrature corresponding to the uncertainty of the Cs fountain itself, the time transfer uncertainty in comparing the Cs fountain to TAI, and the NIST timescale to TAI, and a relatively negligible term arising from comparing maser S3 to the NIST timescale[13,29,30]. There is clear evidence of component ageing related drift for this maser and this behaviour is typical for all four masers considered. In this figure, the fractional frequency differences are suppressed by a factor of 100 so that the estimated uncertainties and the fractional frequency differences can be shown in the same plot. Only 20% of the data are plotted to avoid blotting out the curves (for more information, see Supplementary Table 2). One of the reasons this study is an improvement over the previous effort is that there are significantly more fountain data after MJD 54000 (ref. [6]). The identifiers used for some of the Cs fountains or the host laboratories may have changed over the years.

this aspect of LPI that takes one of the postulates applicable to metric theories of gravitation closer to general relativity through an additional requirement: the principle of general covariance. It states that the laws of physics ought to be expressible in a coordinate-independent formalism; constants of nature comprise one part of that story. In the following paragraphs, we use the result of equation (3) together with the previous best estimates of β from Guéna et al. (2012)[15] and Peil et al. (2013)[16] to place constraints on the variations of two fundamental constants that are related to hyperfine transitions.

Matter and its interaction with fields may be parameterized in terms of the masses of quarks, the mass of an electron, the fine structure constant α, and the quantum chromodynamics (QCD) energy scale parameter Λ_{QCD} at which the QCD coupling begins to diverge[17]. α describes electromagnetic interactions in matter and Λ_{QCD} measures the strong interaction. For example, $\alpha\sim1/137$, can be interpreted as the ratio of the speed of an electron in the Bohr atom to the speed of light (the photon is the force carrier for the electromagnetic force) in vacuum. Dicke, in his 1963 lectures, conjectured that $c^2(\mathrm{d}\alpha/\mathrm{d}t)\approx\alpha^2(\mathrm{d}\Phi/\mathrm{d}t)$, where $\mathrm{d}\Phi/\mathrm{d}t$ is the annual variation of the gravitational potential due to the Sun[4].

The difference in frequency shifts due to hyperfine splitting for a pair of clocks may be recast as a variation of the ratio of the frequencies, which is related to the variation of the fundamental constants by the formula[15,18–20]

$$\mathrm{dln}\left(f_{\mathrm{A}}/f_{\mathrm{B}}\right)=\Delta K_{\alpha}\mathrm{dln}\alpha+\Delta K_{\mathrm{q}}\mathrm{dln}X_{\mathrm{q}} \tag{4}$$

where α is the fine structure constant and $X_{\mathrm{q}}=m_{\mathrm{q}}/\Lambda_{\mathrm{QCD}}$ is the ratio of the light quark mass to the QCD scale. K_{α} and K_{q} are the relative sensitivities of the hyperfine relativistic factor and nuclear magnetic moment to the variations of α and X_{q} respectively. Since the H masers used in this study are susceptible to drifts whose origins are not well understood, over periods that are of the order of a few years, below we present a formalism to constrain $\mathrm{dln}(f_{\mathrm{H}}/f_{\mathrm{Cs}})$.

The ratio of the hyperfine frequencies of two atomic species is related to the spatial variation of gravitational potential, from equation (2), which can also be written as:

$$\mathrm{dln}\left(f_{\mathrm{H}}/f_{\mathrm{Cs}}\right)=(\beta\Delta\Phi)/c^2 \tag{5}$$

where c, β and $\Delta\Phi$ are the same quantities as in equations (1) and (2). To constrain α and X_{q} individually, first we note that

$$\delta\alpha/\alpha=k_{\alpha}\delta\left(\Phi/c^2\right)\ \text{and}\ \delta X_{\mathrm{q}}/X_{\mathrm{q}}=k_{\mathrm{q}}\delta\left(\Phi/c^2\right) \tag{6}$$

where k_{α} and k_{q} are dimensionless coupling constants linking the variation of α and X_{q} to the variation of the gravitational potential, with δ denoting the variation of fundamental constants and gravitational potential. Using equation (6) in equation (4) and rearranging the terms, we obtain equations of the form

$$\beta|_{\mathrm{H-Cs}}=\Delta K_{\alpha}k_{\alpha}+\Delta K_{\mathrm{q}}k_{\mathrm{q}} \tag{7}$$

We may use the previous best estimates for β involving clock transitions that depend on the hyperfine splitting analysed in this study to solve for the dimensionless coupling constants (see Table 1).

Using the entries of Table 1 in equation (7) yields two independent sets of values for k_{α} and k_{q} for equations involving pairs (i) and (iii), and (ii) and (iii) of Table 1. The equally weighted averages of the two values for both k_{α} and k_{q} yield:

$$k_{\alpha}=(0.70\pm1.8)\times10^{-7}\ \text{and}\ k_{\mathrm{q}}=(-25\pm21)\times10^{-7} \tag{8}$$

The previous best estimate for $k_{\mathrm{q}}=(3.8\pm4.9)\times10^{-6}$ was reported by Peil at al. (2013)[16]. More recently, Dzuba and Flambaum (2017) report a slightly better value of $k_{\alpha}=(-0.53\pm1.0)\times10^{-7}$ (ref. [21]). Our results are an improvement over the previous estimates of k_{q} by a factor of two. The combined annual variation of gravitational potential due to the Sun and Jupiter based on the ephemerides is 3.313×10^{-10}. Using this value in equation (6)

$$\begin{aligned}\dot{\alpha}/\alpha &= (2.3\pm6.0)\times10^{-17}\ \mathrm{yr}^{-1}\ \text{and}\\ \dot{X}_{\mathrm{q}}/X_{\mathrm{q}} &= (-8.3\pm7.0)\times10^{-16}\ \mathrm{yr}^{-1}\end{aligned} \tag{9}$$

Godun et al. (2014)[22] estimated $\dot{\alpha}/\alpha=(-0.7\pm2.1)\times10^{-17}\,\mathrm{yr}^{-1}$ from direct measurements of clock frequency drifts—a factor of three better than the results presented here that are inferred from the estimated value of β and the relative sensitivities (see Table 1). Similar measurements by Guéna et al. (2012)[15] had set the previous

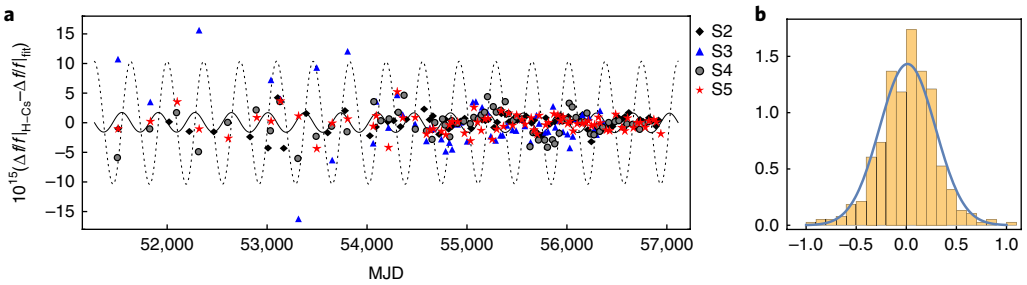

Fig. 3 | Addressing the drift in the H maser with fitting functions. A third-order polynomial was used to fit the fractional frequency difference data to phenomenologically account for component ageing and any other systematic shifts that can be as high as $\Delta f/f \approx 10^{-14}$ per year. This should not affect the correlation sought between the residuals of the fit and the sinusoidal solar potential variation. The data are roughly split into three segments of five years each, such that the residuals from the fit conform to a normal distribution. Data segmentation was necessitated by the requirement to keep the number of free parameters in the fit to a minimum and to keep the fitting functions simple. Only 20% of the data points are shown to avoid blotting out the curves. **a**, Residuals after fitting the fractional frequency difference data with third-order polynomials. The solid and dotted curves are the gravitational potential variation and its time derivative (with different units) as described in Fig. 2. S3 has a larger scatter compared to other masers. **b**, Histogram of the 380 residuals for maser S2—as an example—scaled to a normal probability distribution function (vertical axis). The horizontal axis shows the residuals from the third-order polynomial fit scaled to unity so that the probability distribution function (solid curve) and the data can be presented in the same plot. The actual values range between about $\pm 5.0 \times 10^{-15}$.

Table 1 | Comparison of the previous best estimates for β with the results of the present study

No.	Reference	$\beta_A - \beta_B$	A, B	ΔK_α	ΔK_q
(i)	Guéna et al., 2012[15]	$(0.11 \pm 1.04) \times 10^{-6}$	Rb, Cs	−0.49	−0.025
(ii)	Peil et al., 2013[16]	$(-2.7 \pm 4.9) \times 10^{-7}$	Rb, H	0.34	0.084
(iii)	This work	$(2.24 \pm 2.48) \times 10^{-7}$	H, Cs	−0.83	−0.110

The values for the differences in K_α and K_q are from Flambaum and Tedesco, 2006[19].

best estimates for $\dot{X}_q / X_q = (0.14 \pm 9.0) \times 10^{-16}\,\mathrm{yr}^{-1}$, as inferred by Huntemann et al. (2014)[23].

Since LPI—as a postulate of general relativity—is more general than any experiment involving only two atomic species, combining the values of LPI parameters from Table 1, we obtain the weighted average

$$\beta = (2.2 \pm 2.2) \times 10^{-7} \quad (10)$$

with assigned weights that are equal to the inverse of the square of the uncertainties. A null hypothesis ($\beta = 0$) is a necessary condition for general relativity to be valid, but no experiment can serve as a sufficient condition, since all experiments have finite errors[4]. By deriving new limits on the variations of two fundamental constants, we were able to extend the applicability of the null hypothesis of LPI for validating metric theories, which are a more general class of theories, to general relativity. The implications of varying fundamental constants in the context of unified theories and alternatives to general relativity are detailed in Uzan (2003)[17].

We note that using three masers instead of four made only a small difference in the estimation of β. The use of more data is unlikely to yield stricter constraints. Owing to the long-term drifts that are typical in H masers, there is not much likelihood of improving the uncertainty in the LPI parameter using H masers and Cs fountains. Future improvements may come from comparisons of optical clocks, which might perform at least two orders of magnitude better—only limited by the uncertainty in the estimation of the total gravitational potential variation—than comparisons between H masers and Cs-fountain standards because the performance of optical clocks continues to improve as noise in such clocks is better understood[24].

Of the many challenges in comparing different optical clock types, up until recently, the main ones have been the availability of frequency links with stability and frequency transfer uncertainty comparable to the best optical clocks, and availability of robust clocks capable of running simultaneously over periods that match or exceed Earth's orbital period. An example of the improvement in the development of fibre links is the recently commissioned 1,415 km telecom fibre link connecting Paris and Braunschweig[25]. Work is also underway to compare the NIST ytterbium clock and the JILA strontium clock using a fibre link[26,27]. These optical clocks and fibre links are two important aspects of any future experiments that are certain to improve the results presented in this paper, at which time the gravitational perturbations from Jupiter will not be negligible as they were for this study[28].

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41567-018-0156-2.

Received: 30 May 2017; Accepted: 25 April 2018;
Published online: 4 June 2018

References

1. Einstein, A. *The Collected Papers of Albert Einstein, Volume 6: The Berlin Years: Writings, 1914–1917* (Princeton Univ. Press, Princeton, NJ, 1996).
2. DeMille, D., Doyle, J. M. & Sushkov, A. O. Probing the frontiers of particle physics with tabletop-scale experiments. *Science* **357**, 990–994 (2017).
3. Will, C. M. The confrontation between general relativity and experiment. *Living Rev. Relativ* **9**, 3 (2014).
4. Dicke, R. H. *Experimental Relativity: Lectures delivered at Les Houches, 1963; Relativity, Groups and Topology* (eds DeWitt, C. and DeWitt, B.) (Gordon & Breach, New York, NY, 1964).
5. Sullivan, D. B. Time and frequency measurement at NIST: the first 100 years. In *Proc. 2001 IEEE International Frequency Control Symposium*, Cat. No. 01CH37218, 4–17 (IEEE, Piscataway, NJ, 2001).
6. Ashby, N. et al. Testing local position invariance with four cesium-fountain primary frequency standards and four NIST hydrogen masers. *Phys. Rev. Lett.* **98**, 070802 (2007).
7. Levi, F. et al. Accuracy evaluation of ITCsF2: a nitrogen cooled caesium fountain. *Metrologia* **51**, 270–284 (2014).
8. Jefferts, S. R. et al. Accuracy evaluation of NIST-F1. *Metrologia* **39**, 321–336 (2003).
9. Weyers, S., Hubner, U., Schroder, R., Tamm, C. & Bauch, A. Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB. *Metrologia* **38**, 343–352 (2001).

10. Gerginov, V. et al. Uncertainty evaluation of the caesium fountain clock PTB-CSF2. *Metrologia* **47**, 65–79 (2010).
11. Szymaniec, K. et al. NPL Cs fountain frequency standards and the quest for the ultimate accuracy. *J. Phys. Conf. Ser.* **723**, 012003 (2016).
12. Guéna, J. et al. Progress in atomic fountains at LNE-SYRTE. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **59**, 391–410 (2012).
13. *Reports of Evaluation of Primary Frequency Standards* (BIPM, accessed 14 March 2017); http://www.bipm.org/en/bipm-services/timescales/time-ftp/data.html
14. Folkner, W. M., Williams, J. G., Boggs, D. H., Park, R. S. & Kuchynka, P. *The Planetary and Lunar Ephemerides DE430 and DE431* InterPlanetary Network Progress Report 42–196 (IPN, 2014).
15. Guéna, J. et al. Improved tests of local position invariance using Rb87 and Cs133 fountains. *Phys. Rev. Lett.* **109**, 080801 (2012).
16. Peil, S., Crane, S., Hanssen, J. L., Swanson, T. B. & Ekstrom, C. R. Tests of local position invariance using continuously running atomic clocks. *Phys. Rev. A* **87**, 010102(R) (2013).
17. Uzan, J. The fundamental constants and their variation: observational and theoretical status. *Rev. Mod. Phys.* **75**, 403–455 (2003).
18. Fischer, M. et al. New limits on the drift of fundamental constants from laboratory measurements. *Phys. Rev. Lett.* **92**, 230802 (2004).
19. Flambaum, V. V. & Tedesco, A. F. Dependence of nuclear magnetic moments on quark masses and limits on temporal variation of fundamental constants from atomic clock experiments. *Phys. Rev. C* **73**, 055501 (2006).
20. Dinh, T. H., Dunning, A., Dzuba, V. A. & Flambaum, V. V. Sensitivity of hyperfine structure to nuclear radius and quark mass variation. *Phys. Rev. A* **79**, 054102 (2009).
21. Dzuba, V. A. & Flambaum, V. V. Limits on gravitational Einstein equivalence principle violation from monitoring atomic clock frequencies during a year. *Phys. Rev. D* **95**, 015019 (2017).
22. Godun, R. M. et al. Frequency ratio of two optical clock transitions in $^{171}Yb^{+}$ and constraints on the time variation of fundamental constants. *Phys. Rev. Lett.* **113**, 210801 (2014).
23. Huntemann, N. et al. Improved limit on a temporal variation of m_p/m_e from comparisons of Yb^{+} and Cs atomic clocks. *Phys. Rev. Lett.* **113**, 210802 (2014).
24. Ludlow, A. D., Boyd, M. M., Ye, J., Peik, E. & Schmidt, P. O. Optical atomic clocks. *Rev. Mod. Phys.* **87**, 637–701 (2015).
25. Lisdat, C. et al. A clock network for geodesy and fundamental science. *Nat. Commun.* **7**, 12443 (2016).
26. Nicholson, T. L. et al. Systematic evaluation of an atomic clock at 2×10^{-18} total uncertainty. *Nat. Commun.* **6**, 6896 (2015).
27. Schioppo, M. et al. Ultrastable optical clock with two cold-atom ensembles. *Nat. Photon.* **11**, 48–52 (2017).
28. Tobar, M. E. et al. Testing local position and fundamental constant invariance due to periodic gravitational and boost using long-term comparison of the SYRTE atomic fountains and H-masers. *Phys. Rev. D* **87**, 122004 (2013).
29. Parker, T. E., Jefferts, S. R., Heavner, T. P. & Donley, E. A. Operation of the NIST-F1 caesium fountain primary frequency standard with a maser ensemble, including the impact of frequency transfer noise. *Metrologia* **42**, 423–430 (2005).
30. Panfilo, G. & Parker, T. E. A theoretical and experimental analysis of frequency transfer uncertainty, including frequency transfer into TAI. *Metrologia* **47**, 552–560 (2010).

Acknowledgements

We acknowledge funding from NASA grant NNH12AT81I. We also thank the atomic standards group at NIST for maintaining the H masers and sharing the data. We thank E. Donley, S. Jefferts and C. Oates for providing valuable suggestions that have helped improve this paper. We thank J. Sherman and J. Ye for discussing the planned clock comparisons between NIST and JILA. This work is a contribution of NIST and is not subject to US copyright.

Author contributions

T.E.P. compiled data from BIPM and analysed maser versus primary frequency standard data. N.A. and B.R.P. carried out the data analysis and performed the calculations that were the basis of the main conclusions of this paper. B.R.P. wrote the manuscript with input from all of the authors, and all authors discussed the results and the conclusions.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41567-018-0156-2.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to B.R.P.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Background and scope. In 1978, Turneaure et al. compared two H masers with a set of three superconducting cavity-stabilized oscillators as the solar gravitational potential changed due to Earth rotation[31]. Measurements over a ten-day period were consistent with LPI and EEP at about the 2% level. Godone et al.[32] compared Mg and Cs standards for 430 days and were able to improve on the result of Turneaure et al. by a factor of almost 20. In 2012, Guena and co-workers at SYRTE were able to compare Cs and Rb laser-cooled atomic fountain clocks over a period of 14 years, using variations in the solar gravitational potential to place significant limits on the rate of change of the fractional frequency difference of the two clocks, to obtain $\beta|_{\mathrm{Rb-Cs}} = (1.1 \pm 10.4) \times 10^{-7}$[15]. In 2013, Peil and co-workers at the United States Naval Observatory used continually running clocks (Rb fountains and H masers) for 1.5 years and reported a value of $\beta|_{\mathrm{Rb-H}} = (-2.7 \pm 4.9) \times 10^{-7}$ (ref. [16]). For comparison with this, we quote the result of previous comparisons at NIST for H and Cs $\beta|_{\mathrm{H-Cs}} = (1.0 \pm 14) \times 10^{-7}$ (ref. [6]).

The advances reported here in testing LPI are complementary to at least one planned space-based experiment for testing the postulates of metric theories of gravity. The Atomic Clock Ensemble in Space project, comprising a H maser and a Cs tube on-board the International Space Station and microwave links for comparing with ground-based clocks, will aim to test the gravitational redshift and LLI[33]. The now called-off, but nevertheless highly rated science experiment, Space-Time Explorer and Quantum Equivalence Principle Space Test had plans to test the WEP using atom interferometry[34]. As clocks become more portable and space-qualified, one could foresee more such experiments planned well into the future.

Gravitational redshift versus LPI. In an accelerated laboratory, if two otherwise identical clocks separated by height h exchange photons, the photon frequencies will suffer first-order Doppler shifts due to the velocity difference that builds up during the propagation delay between clocks, because the speed of light is finite. This implies clocks at different gravitational potentials will suffer frequency shifts that do not depend on the structure of the clocks. A comparison of the frequencies of two similar clocks at different locations can be considered as a nonlocal gravitational experiment and understood within the framework of EEP. The gravitational redshift described above has been measured accurately to 120 parts per million[35].

Correcting for H-maser drift. Environmental factors affecting H masers are studied in detail in Parker (1999)[36] and the impact of frequency transfer noise in comparing masers and Cs fountains is described in Parker et al. (2005)[29] (see also, ref. [6]). Since all of the masers are housed in the same location, the stated uncertainty is the same for all of the masers; the corrections for frequency fluctuations due to changes in environmental factors are different. In addition to environmental factors, the masers experience long-term drifts that are related to component ageing[36,37]. Frequency shifts for all H masers are referenced to the location of NIST, Boulder.

Data analysis for estimating β. A more detailed procedure for obtaining β is described below. For each maser, optimizing equation (2)

$$\sum_i \frac{\mathrm{d}}{\mathrm{d}\beta}(\Delta_i - \beta\Delta\Phi_i / c^2)^2 = 0 \tag{11}$$

where $\Delta_i = \Delta f/f|_{\mathrm{H-Cs}} - \Delta f/f|_{\mathrm{fit}}$, yields

$$\beta_j = c^2 \frac{\sum_i \Delta_i \Delta\Phi_i}{\sum_i \Delta\Phi_i^2} \tag{12}$$

where the index i is the time stamp label for a data point. For example, i varies from 1 to 380 for maser S2 ($j=1$) (see Supplementary Table 1). $\Delta\Phi$ is the change in total gravitational potential. The maximum and minimum values of frequency difference for each maser versus fountain after correcting for environmental effects are also given in Supplementary Table 1. Maser frequency drifts are analysed and quantified in Ashby et al. (2007)[6]. The uncertainty for an individual maser correlated with the gravitational potential is

$$\langle \delta\beta_j^2 \rangle = c^4 \frac{\left\langle \sum_{ik} \delta(\Delta_i)\Delta\Phi_i \delta(\Delta_k)\Delta\Phi_i \right\rangle}{\left(\sum_i \Delta\Phi_i^2\right)^2} \tag{13}$$

where $\langle \delta(\Delta_i)\delta(\Delta_k)\rangle = \delta_{ik}\delta(\Delta_i)^2$, to give

$$\delta\beta_j = c^2 \frac{\left(\sum_i \delta(\Delta_i)^2 \Delta\Phi_i^2\right)^{1/2}}{\sum_i \Delta\Phi_i^2} \tag{14}$$

where $\delta(\Delta_i)$ is the combined uncertainty (see Fig. 2). There has been significant improvement in the reported uncertainties in the last seven years of the data set compared to the first seven years (before MJD 54000) (see Supplementary Fig. 1).

The final result is obtained by taking the weighted average and adding the uncertainties in quadrature

$$\beta = \sum_j w_j \beta_j = \left(\sum_j \frac{1}{\delta\beta_j^2}\right)^{-1} \left(\sum_j \frac{\beta_j}{\delta\beta_j^2}\right) \tag{15}$$

where w_j are the weights. The 1σ uncertainty is obtained by deriving a probability distribution function for a normal distribution for the residuals (see Fig. 3), from which β_j is obtained. We provide the final result for the probability distribution function

$$\mathcal{P}\left(\beta = \sum w_j \beta_j\right) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sum_j w_j^2 \delta\beta_j^2}\right)^{1/2} \exp\left(-\frac{\left(\beta - \sum_j w_j \delta\beta_j\right)^2}{2\sum_j w_j^2 \delta\beta_j^2}\right) \tag{16}$$

The combined 1σ uncertainty is $\delta\beta = \sqrt{\sum_j w_j^2 \delta\beta_j^2}$.

Data availability. The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

References

31. Turneaure, J. P., Will, C. M., Farrell, B. F., Mattison, E. M. & Vessot, R. F. C. Test of the principle of equivalence by a null gravitational red-shift experiment. *Phys. Rev. D* **27**, 1705–1714 (1983).
32. Godone, A., Novero, C. & Tavella, P. Null gravitational redshift experiment with nonidentical atomic clocks. *Phys. Rev. D* **51**, 319–323 (1995).
33. Heß, M. P. et al. The ACES mission: system development and test status. *Acta Astronaut.* **69**, 929–938 (2011).
34. Altschul, B. et al. Quantum tests of the Einstein equivalence principle with the STE-QUEST space mission. *Adv. Space Res.* **55**, 501–524 (2015).
35. Vessot, R. F. C. et al. Test of relativistic gravitation with a space-borne hydrogen maser. *Phys. Rev. Lett.* **45**, 2081–2084 (1980).
36. Parker, T. E. Environmental factors and hydrogen maser frequency stability. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **46**, 745–751 (1999).
37. Lewis, L. L. An introduction to frequency standards. *Proc. IEEE* **79**, 927–935 (1991).