Metrologia

PAPER

Atomic spectroscopy for primary thermometry

To cite this article: G-W Truong *et al* 2015 *Metrologia* **52** S324

IOP Publishing | Bureau International des Poids et Mesures Metrologia

Metrologia **52** (2015) S324–S342 doi:10.1088/0026-1394/52/5/S324

Atomic spectroscopy for primary thermometry

G-W Truong[1,2], D Stuart[3], J D Anstie[1], E F May[4], T M Stace[5] and A N Luiten[1,2]

[1] Institute for Photonics and Advanced Sensing (IPAS) and School of Chemistry and Physics, The University of Adelaide, Adelaide, SA 5005, Australia
[2] School of Physics, The University of Western Australia, Perth, Western Australia 6009, Australia
[3] School of Physics, The University of Oxford, UK
[4] Centre for Energy, School of Mechanical and Chemical Engineering, The University of Western Australia, Crawley, Western Australia 6009, Australia
[5] ARC Centre for Engineered Quantum Systems, University of Queensland, Brisbane 4072, Australia

E-mail: andre.luiten@adelaide.edu.au

Received 28 February 2015, revised 25 May 2015
Accepted for publication 27 May 2015
Published 19 August 2015

Abstract

Spectroscopy has been a key driver and motivator of new understanding at the heart of physics. Here we describe high-precision measurements of the absorption lineshape of an atomic gas with an aim towards primary thermometry. We describe our progress in pushing this type of spectroscopy to the ultimate limit, in particular in describing experimental work with Rubidium and Cesium, although we also consider the potential for other elements in expanding the precision, accuracy and range of the approach. We describe the important technical and theoretical limits which need to be overcome in order to obtain accurate and precise results—these challenges are not unique to atomic spectroscopy but are likely to afflict all high precision spectroscopy measurements. We obtain a value for $k_B = 1.380\,545(98) \times 10^{-23}$ J K^{-1} where the 71 ppm uncertainty arises with difficulties in defining the Lorentzian component of the lineshape.

Keywords: Boltzmann constant, laser spectroscopy, atomic spectroscopy, fundamental constants, shot-noise limits, accurate lineshape measurement

(Some figures may appear in colour only in the online journal)

1. Introduction

We present here an overview of our experimental work aimed at high precision measurements of the spectroscopic parameters of atomic vapours. The principal motivation was a desire to develop new techniques to re-measure Boltzmann's constant, k_B, in preparation for a redefinition of the kelvin [1]; Nonetheless, it is clear that advances in absorption lineshape measurement and theory have broader application across many important tasks including thermometry of plasma in tokamaks [2], gas detection [3] and astrophysical observations [4, 5]. Moreover, the accurate measurement of the natural linewidth and transition frequencies is crucial for validating theoretical atomic physics calculations [6–8].

At thermal equilibrium, the velocity distribution of atomic absorbers is related to the temperature through the Boltzmann distribution. This simple and fundamental relationship forms an excellent foundation for a type of primary thermometry known as Doppler broadening thermometry (DBT) [9, 10]. A highly precise measurement of the transmission lineshape of an atomic vapour will deliver information about the velocity distribution and hence measure the temperature of the sample. In this work we will present experimental work on the transmission lineshapes of Rubidium (Rb) and cesium (Cs) vapor, culminating in a quantum-limited transmission uncertainty of 2 ppm in a 1 s measurement with our second-generation spectrometer. This extreme precision allows us to directly detect subtle lineshape perturbations that have not been previously

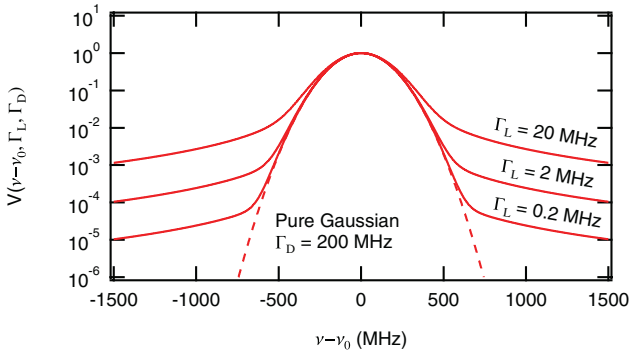

Figure 1. A plot of a Voigt function with fixed Γ_D and several Γ_L.

observed, and which need to be explained before the measurements can be used to provide an accurate measurement of temperature. We thus see how this high precision experimental work motivated the development of a new theoretical model to give a full description of the observations: in turn, this new theory then motivated new experiments to search for new effects predicted by the theory [11]. This increased understanding will also be used to guide us in making an optimal choice of atomic absorber for Doppler broadening thermometry.

The leading technique for primary thermometry uses the speed of sound in a noble gas (as is described in many other articles in this special issue and elsewhere [12–14]). The high performance of that technique sets a challenging target for any competing primary thermometry approach at the ∼ 1 ppm level. Nevertheless, given that there is a great difference of approach between DBT and the primary technique, it is likely that the systematic errors of each are also quite different. Accordingly, there is great deal of value in pushing this new technique as long as it offers a performance in the sub-10 ppm range [15].

2. Choice of quantum absorber

The DBT approach has a strong and transparent theoretical foundation. The spectrum of a gas sample is measured with high precision, and we fit the measured data to that theoretical model. If the model includes all of the relevant physics then one can extract the contribution to spectral broadening arising solely from the thermal distribution of atomic/molecular speeds. Ammonia probed by a frequency-stabilized CO_2 laser at 1.34 μm was the first thermometric substance employed in a deliberate DBT experiment [16]. Subsequently, an extended-cavity diode laser at 2 μm was used to probe a ro-vibrational transition of CO_2 [17]. In these first experiments, the line-shape was assumed to be a Gaussian or a Voigt profile (a Gaussian lineshape convolved with an underlying Lorentzian lineshape). Since 2007, with the ambitious goal of approaching 1 ppm accuracy, substantial experimental and theoretical improvements have been made to DBT using ammonia [18, 19], oxygen [20], ethyne [21], and water [22]. However, one key challenge peculiar to molecular absorbers is the need to account for complex collisional effects on the line shape [19]. Our approach avoids this by using a dilute atomic vapour [23, 24] with a strong dipole transition for which a tractable, microscopic theory has been developed [25, 26]. This section will describe the basic principles that determine a preference for a particular atomic absorber and transition.

The observed absorption line-shape, $\mathcal{T}_{at}(\nu)$, of an atomic transition with a rest-frame transition frequency ν_0 is broadened through two components: a homogenous component that reflects the natural broadening of every atom together with an inhomogeneous contribution that reflects the particular velocity of the absorber in the laboratory frame.

The homogenous broadening has a Lorentzian spectral signature, $L(\nu;\Gamma)$, that can be written as:

$$L(\nu;\Gamma_L) = \frac{1}{1+\left(\frac{\nu-\nu_0}{\Gamma_L}\right)^2} \tag{1}$$

with a natural (half) linewidth $\Gamma_L = 1/(4\pi\tau)$ for a finite excited state lifetime τ. The atomic velocity distribution gives a Gaussian spectral shape that can be written as:

$$G(\nu;\Gamma_D) = \frac{1}{\sqrt{\pi}\Gamma_D} e^{-\left(\frac{\nu-\nu 0}{\Gamma_D}\right)^2} \tag{2}$$

with an e-fold width of $\Gamma_D = \nu_0\sqrt{2k_BT/(mc^2)}$.

The observed spectral shape is a convolution of these two components, termed a Voigt lineshape (see figure 1):

$$V(\nu;\Gamma_L,\Gamma_D) = \frac{\alpha}{C}\int_{-\infty}^{+\infty} d\nu' G(\nu-\nu',\Gamma_D)L(\nu',\Gamma_L) \tag{3}$$

where the normalisation constant C is chosen so that $V(0;\Gamma_L,\Gamma_D) = 1$.

An ideal experiment would utilise a simple two-level atomic system that allows only a single transition; unfortunately, nature has denied this and all real atoms have degeneracies and almost all exhibit multiple overlapping transitions. Thus in any real observation we see an absorption function that is composed of several overlapping transitions. We can now explicitly write Beer's law including the functional dependence of the absorption coefficient and multiple transitions labeled by i as:

$$\mathcal{T}_{at}(\nu) = e^{-\sum_i \alpha_i z V(\nu-\nu_i;\Gamma_L,\Gamma_D)} \tag{4}$$

where z is the length of the absorption cell, $\mathcal{T}_{at}(\nu)$ is the transmission through the cell, and α_i is the absorption coefficient of the ith allowed transition. On the lower part of figure 2 we show the transmission spectra ($\mathcal{T}_{at}$) for the D2 transition of Rb-87 and the D1 transition of Cs-133 for a cell length of $z = 10$ cm. On the upper part of figure 2 we show the individual components, $\alpha_i z V(\nu-\nu_i;\Gamma,\nu_D)$, that contribute to produce the total absorption curve. We note that for the Rb D2 transition there is a strong overlap which means that neither transition is fully resolved.

The simplest atom we can choose is one with a transition with minimal hyperfine structure and a simple Zeeman structure. With this in mind we highlight the key factors in choosing the most ideal quantum absorber in an approximate hierarchy of importance:

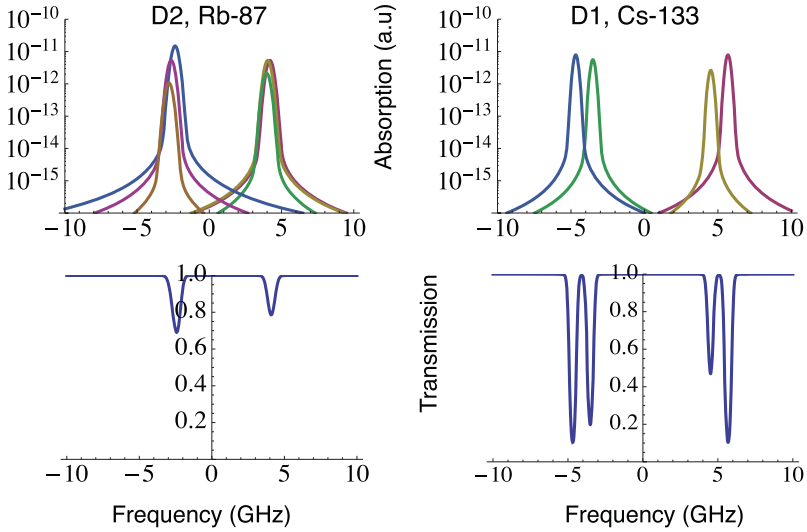

Figure 2. Theoretical Absorption Spectra and Transmission Spectra for Rb-87 and Cs-133 calculated using the ADM package [27].

Table 1. Properties of Alkali D2 Lines: columns show element; resonance transition wavelength; Doppler broadened linewidth; natural linewidth; nuclear moment; absorption coefficient at room temperature; temperature required to obtain a vapour pressure that generates an absorption coefficient of 1 cm^{-1}; ratio of natural to Doppler linewidth; ratio of hyperfine splitting to Doppler linewidth.

Element	λ(nm)	Doppler (MHz)	Natural (MHz)	I	$\alpha_{298K}(cm^{-1})$	T_l(K)	Hyperfine (MHz)	Nat./Dop.	HFS/Dop.
^{133}Cs	894.6	359	4.6	$\frac{7}{2}$	1.9	291	1168	0.013	3.20
^{85}Rb	795.0	506	5.7	$\frac{5}{2}$	2.7×10^{-1}	311	362	0.011	0.72
^{87}Rb	795.0	500	5.7	$\frac{3}{2}$	1.1×10^{-1}	320	812	0.011	1.60
^{39}K	770.1	771	6.0	$\frac{3}{2}$	1.1×10^{-2}	343	58	0.008	0.07
^{23}Na	589.8	1311	9.8	$\frac{3}{2}$	8.8×10^{-6}	416	189	0.007	0.14
^{7}Li	671.0	2086	5.9	$\frac{3}{2}$	7.7×10^{-15}	633	92	0.003	0.04

Note: We see that Cs is the best choice because of its strong transition, resolved hyperfine structure, good vapour pressure and convenient wavelength although the D1 line would be a better choice—see table 2.

- A strong absorber so that the atomic cell is compact (minimal thermal gradients) and can be filled with a vapour that is dilute at a convenient temperature (avoids collisions),
- An optical transition at a convenient wavelength for which high quality laser light is available
- A simple resolvable optical transition (that is all available transitions should have frequency separations either much larger or much smaller than the Doppler broadening)
- A high degree of separability between the homogeneous and inhomogeneous broadening processes. In practice this means that one would like a long excited state lifetime (i.e. $\Gamma_L \ll \Gamma_D$ or alternatively that $\tau \gg c/\nu_0\sqrt{\frac{m}{32\pi^2 k_B T}}$) as well as a value for Γ_L that is extremely well known.
- A low mass atom and/or a high frequency transition so that Γ_D is broad. This allows the use of a probing light source with moderate spectral purity.

Unfortunately, not all of these desirable features are solely beneficial. For example, the first key factor listed above necessarily leads to strong optical pumping effects which complicate the interpretation of the spectra (see below).

We have made an examination of the key features of potential atomic absorbers in order to match them to the requirements listed above; these have been summerized in tables 1–4. The expected spectral features of alkali metals probed on the D2 transition are shown on table 1, with the D1 transition in table 2. The first three elements listed show a sufficient absorption in a convenient length cell near room temperature while almost all (except Na) show a transition with a convenient wavelength for low-cost laser technology. The hyperfine splitting of the excited state (which is always smaller than the hyperfine splitting in the ground state) is seen to be less than the Doppler width in almost every case, but not much smaller. This causes substantial difficulty in fitting the measured data to a model of the form in equation (4) due to overlapping absorption features. An example of this problem is shown on the left-hand side of figure 2—where we simulate the allowed D2 transition in Rb-87 and observe the overlap between the various contributions. The two exceptions that do not have show this overlap between the Doppler broadening and the hyperfine splitting in the upper state are the D1 transition of Rb-87 (marginal) and Cs-133 (better). Finally, the table shows

Table 2. Properties of Alkali D1 Lines: columns show the element, resonance transition wavelength; Doppler broadened linewidth; natural linewidth; nuclear moment, absorption coefficient at room temperature; temperature required to obtain a vapour pressure that generates an absorption coefficient of 1 cm^{-1}; hyperfine splitting of the excited state, ratio of natural to Doppler linewidth and ratio of hyperfine splitting to Doppler linewidth.

Element	λ(nm)	Doppler (MHz)	Natural (MHz)	I	$\alpha_{298K}(cm^{-1})$	T_i(K)	Hyperfine (MHz)	Nat./Dop.	HFS/Dop.
^{133}Cs	852.3	377	5.2	$\frac{7}{2}$	1.9	292	151	0.014	0.40
^{85}Rb	780.2	516	6.1	$\frac{5}{2}$	2.7×10^{-1}	311	50	0.012	0.10
^{87}Rb	780.2	510	6.1	$\frac{3}{2}$	1.1×10^{-1}	321	85	0.012	0.17
^{39}K	766.7	775	6.0	$\frac{3}{2}$	1.1×10^{-2}	343	6	0.008	0.01
^{23}Na	589.2	1312	9.8	$\frac{3}{2}$	8.8×10^{-6}	416	19	0.007	0.01
^{7}Li	671.0	2086	5.9	$\frac{3}{2}$	7.7×10^{-15}	633	−3	0.003	−0.001

Note: We see Cs is the best choice because of its strong transition, resolved hyperfine structure, good vapour pressure and convenient wavelength.

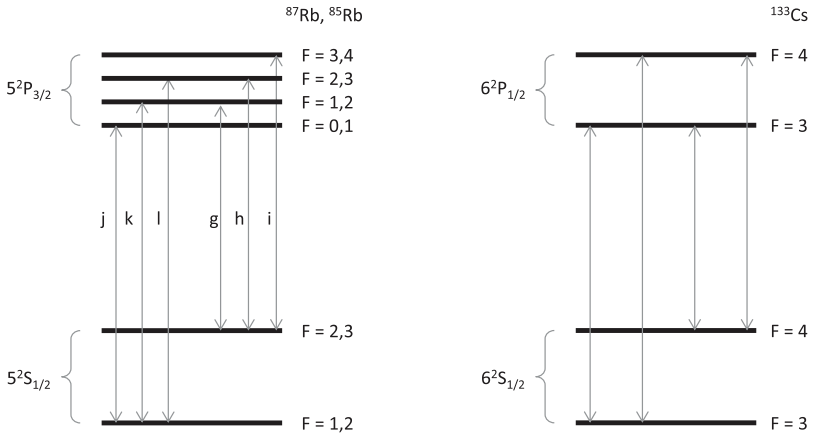

Figure 3. Energy level diagram for D2 transitions in Rb and D1 transition in Cs.

a column $T_i(K)$ that gives the temperature needed to produce a vapour pressure with an absorption length of 1 cm^{-1}—we see that for Na and Li one needs to operate at elevated temperatures while the rest of the alkalis produce reasonable strength absorption in a cell of a few centimetres long at room temperature.

Our first attempt at an atomic DBT system made use of Rb vapour on the D2 transition at 780 nm—this met the first two key requirements listed earlier but violated the third requirement by showing an overlapped spectrum on the upper hyperfine splitting (see figure 2). The energy level diagram for the D2 transition in Rb is shown on the left hand side of figure 3.

Our second generation DBT system was based on the Cs D1 transition (at 895 nm) which satisfies the first three key requirements listed earlier due to the much larger hyperfine splitting in the excited state. The energy level diagram for this transition is shown on the right hand side of figure 3 while the fully resolved spectra is shown on figure 2. Unfortunately, the ratio of Γ_D and Γ_L is not as large as one would hope and this leads to a correlation between Γ_D and Γ_L which are treated as independent parameters in the nonlinear fitting (as will be described below).

In order to best satisfy all the criteria we have considered a wider range of elements: in tables 3 and 4 we consider the singlet and triplet transitions of group II and group II-like elements. We have particularly focussed on those isotopes of these elements that avoid complex hyperfine structures. All elements (except Mercury) require much higher temperatures to enable sufficient optical absorption in a convenient cell size (200–400°C). Conversely, these elements could be important for the development of new high-temperature primary thermometers given that existing approaches have far greater uncertainties at such temperatures than they do near that of the triple point of water. For example, a Ca-40 cell of 1 cm long would provide a structure-free Doppler-broadened line of 1387 MHz width at a temperature of 655 K. For better separation of the natural and doppler widths, one might consider the inter-combination transition lines in the spectra of these atoms rather than the main resonance transitions (see table 4). Although much weaker than the main resonance line, the advantage is that the excited state lifetime can be very long which allows the Lorentzian component to be nicely separated from the Doppler broadening. Table 4 also states what pressure is necessary to present an absorption of 1 cm^{-1}, which will deliver a high signal to noise ratio absorption feature in a short cell. We note that in some of these cases (e.g. Mg) it is likely that the pressure is sufficiently high that collisional effects will contribute to

Table 3. Properties of the Alkali-Earth Singlet Lines ${}^1S_0 - {}^1P_1$: columns in turn show the resonance transition wavelength; the Doppler broadened linewidth; the natural linewidth; the absorption coefficient at room temperature; and the temperature required to obtain an absorption coefficient of 1 cm^{-1}; pressure at temperature T_l; ratio of natural to Doppler linewidth. The use of these isotopes eliminates the hyperfine splitting in the excited state which simplifies the spectra substantially.

Element	λ(nm)	Doppler (MHz)	Natural (MHz)	$\alpha_{298K}(cm^{-1})$	T_l(K)	p_1(mbar)	Nat./Dop.
^{24}Mg	285.3	3585	78	6×10^{-12}	544	8×10^{-6}	3×10^{-2}
^{40}Ca	422.8	2055	35	4×10^{-17}	655	4×10^{-6}	3×10^{-2}
^{88}Sr	460.9	1207	32	4×10^{-14}	590	2×10^{-6}	4×10^{-2}
^{138}Ba	553.7	853	19	2×10^{-17}	667	3×10^{-6}	3×10^{-2}
^{64}Zn	213.9	2753	113	5×10^{-9}	481	1×10^{-5}	5×10^{-2}
^{114}Cd	228.9	1797	84	4×10^{-6}	418	1×10^{-5}	6×10^{-2}
^{174}Yb	398.9	973	31	6×10^{-13}	568	7×10^{-6}	4×10^{-2}
^{202}Hg	184.9	1259	119	4×10^{2}	237	5×10^{-6}	8×10^{-2}

Note: However, the much lower optical cross-section and lower vapour pressures in general means that the operating temperatures need to be much higher to see significant absorption. Mercury is an interesting choice because it can operate at much lower temperatures although the wavelength is inconvenient.

Table 4. Alkali-Earth Triplet Lines ${}^1S_0 - {}^3P_1$: columns in turn show the resonance transition wavelength; the Doppler broadened linewidth; the natural linewidth; the absorption coefficient at room temperature; and the temperature required to obtain a vapour pressure that gives an absorption coefficient of 1 cm^{-1}; pressure at temperature T_l; ratio of natural to Doppler linewidth.

Element	λ(nm)	Doppler (MHz)	Natural (MHz)	$\alpha_{298K}(cm^{-1})$	T_l(K)	p_1(mbar)	Nat./Dop.
^{24}Mg	457.2	2977	4.0×10^{-5}	1×10^{-17}	964	8	2×10^{-8}
^{40}Ca	657.5	1620	4.1×10^{-4}	2×10^{-21}	983	2×10^{-1}	5×10^{-7}
^{88}Sr	689.4	923	7.5×10^{-3}	3×10^{-17}	771	5×10^{-3}	1×10^{-5}
^{138}Ba	791.4	707	1.6×10^{-3}	5×10^{-21}	936	2×10^{-2}	4×10^{-6}
^{64}Zn	307.7	2271	5.2×10^{-3}	6×10^{-13}	677	1×10^{-1}	3×10^{-6}
^{114}Cd	326.2	1411	6.5×10^{-2}	1×10^{-8}	523	7×10^{-3}	6×10^{-5}
^{174}Yb	555.8	751	1.8×10^{-1}	1×10^{-14}	658	5×10^{-4}	4×10^{-4}
^{202}Hg	253.7	979	1.3	1×10^{1}	270	2×10^{-4}	1×10^{-3}

Note: The combination of low optical cross-section and vapour pressure (except for Mercury) means that these vapours could only be used at elevated temperatures although the ratio of natural to Doppler width is very favourable. Mercury poses a very interesting possibility for use near room temperature with favourable properties.

the lineshape and are thus perhaps not the best choice for a DBT experiment.

Perhaps one of the most interesting transitions in meeting the requirements listed above, but which has not been heavily explored in a DBT application, is Mercury probed on its triplet transition (see table 4). The biggest disadvantage here is the inconvenient wavelength of the transition (254 nm) although modern nonlinear and laser technology has been demonstrated to efficiently provide the small amounts of light required for these experiments [28–30]. The big advantage, as can be seen in table 4, is that the ratio of internal and external broadening processes is an order of magnitude more favourable when compared with Cs D1 transition. This alleviates the necessity to have extremely accurate lifetime data for the excited state.

3. Apparatus

We describe here our second-generation spectrometer as pictured in figure 4. This incorporates the various key elements that we found were essential to making a precise measurement of the transmission of a gas cell. In other work we have described a first generation setup [24] for probing a Rb transition. We make use of a rapidly tunable extended-cavity diode laser as the primary probe laser source. Unfortunately, the diode laser emission contains not only the coherent light emission but also some spontaneous emission which contaminates the transmission measurement. Furthermore, the frequency and amplitude stability of the raw diode laser source is insufficient to obtain high quality results and thus needs to be tempered with additional layers of control.

The probe laser is spectrally and spatially filtered using a combination of an optical cavity of moderate finesse ($\mathcal{F}\approx 305$) and single-mode fibre. The resonant frequency of the cavity was actively controlled to ensure that it is tightly locked to the light field. The spectral filtering reduces the spontaneous emission content of the probe beam from 1.6% to below 0.01%. At the output of the cavity was an AOM to actively control the power incident into the atomic cell which was followed by a single mode fibre to suppresses alignment fluctuations which otherwise lead to noise on the photodiodes. This fibre also guaranteed that the incident field is in a well-defined spatial mode.

The probe laser is frequency stabilised by locking it a user-selectable offset frequency to a master laser that is in turn frequency locked to the $6S_{1/2}\ F = 4$ to $6P_{1/2}\ F = 3$ transition of the D1 line in Cs at 894.6054 nm. This master laser has a frequency instability with a square-root-Allan-variance (SRAV) of ~2 kHz for timescales between 1 s and 30 s. These are the

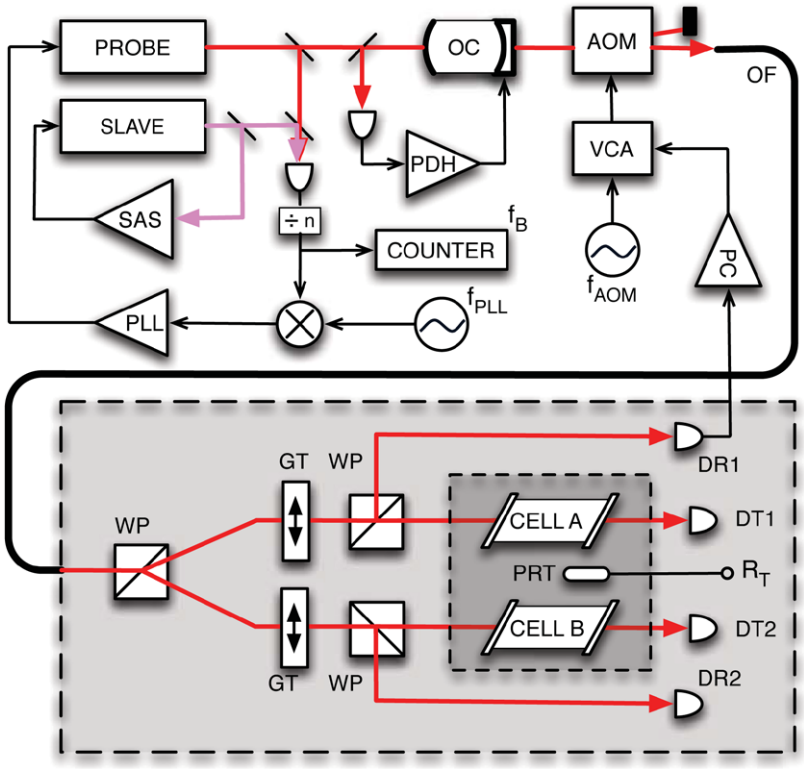

Figure 4. The optical arrangement of the experiment.

relevant time-scales for the spectroscopy, reflecting the time to take a single scan. The probe laser is frequency locked at a user-selectable frequency offset by stabilising the heterodyne beat frequency against a tunable radio frequency (RF) reference oscillator. This frequency lock was sufficiently good that it contributed no additional frequency instability to this slave laser. The RF reference oscillator was step tuned under computer control to ensure that the transmission data were synchronous with the selected frequency. The wavelength of the slave laser was centered on the twin absorption peaks at 894.5793 nm, corresponding to the $6S_{1/2}$ $F = 3$ to $6P_{1/2}$ $F = 3$ and $F = 4$ transitions. The offset frequency was tuned in increments of 50 MHz over a 6500 MHz span that captured the atomic absorption features. Further details on the performance of this probe laser system can be found in the supplementary information in [31].

The light is then delivered into a vacuum chamber in which the vapour cell is mounted in a thermal and magnetic shield (see figure 5). The temperature of the shield can be controlled to a few millikelvin and gradients are suppressed to the same level. This will be described in more detail below. The light is split into two output signals using a combination of a Glan–Taylor polarising prism and a Wollaston beam-splitter. We have shown that the ratio of powers in the output beams is stable to better than 10^{-6} which is not the case when one uses conventional plate or dielectric beam splitters. One output of the beamsplitter directly illuminates a photodiode to give the incident power, whilst the other passes through the vapour cell and is then detected. The ratio of these photodiode signals gives us the transmission ratio, $\mathcal{T}$, which we compare to our theoretical expression,$\mathcal{T}_{at}(\nu)$ given in equation (4).

4. Thermal and magnetic shielding

For the retrieval of an unbiased spectrum, and for long term stability, it is crucial to stabilise the temperature and magnetic environment of the atomic vapour cells. We constructed a thermal isolator consisting of three nested coaxial cylindrical shells in an evacuated chamber (see figure 12) so that heat flow between them is dominated by radiative coupling. Figure 6 shows a schematic view of the shells and the principal heat flows at equilibrium. Two aluminium shells (S1, S2) surround a copper heat reservoir (S3) that is in direct contact with two nominally identical quartz sample cells (C1, C2) with Brewster-angled windows containing elemental cesium (Triad Technologies TT-CS-19X75-Q-BA). A layer of magnetic shielding (B) was present around S3, but due to its much lower mass and conductivity, it was assumed to be irrelevant to the thermal processes. The choice was made to mount two cells in the same thermal environment so as to permit comparisons of spectral parameters across two nominally identical systems.

4.1. Thermal design

A radiative heat load Q impinges on shield S1 from the laboratory environment and is pumped away from the cells and into a chilled baseplate using a thermoelectric cooler (TEC). In this way, the temperature of the cells could be stabilized to a desired set point in the range 273 K–300 K. Shield S1 was designed to have low thermal inertia (low mass and heat capacity) and high thermal diffusivity to permit fast compensation and spatial homogeneity of any laboratory temperature fluctuations and gradients. High diffusivity also helps to minimise the thermal

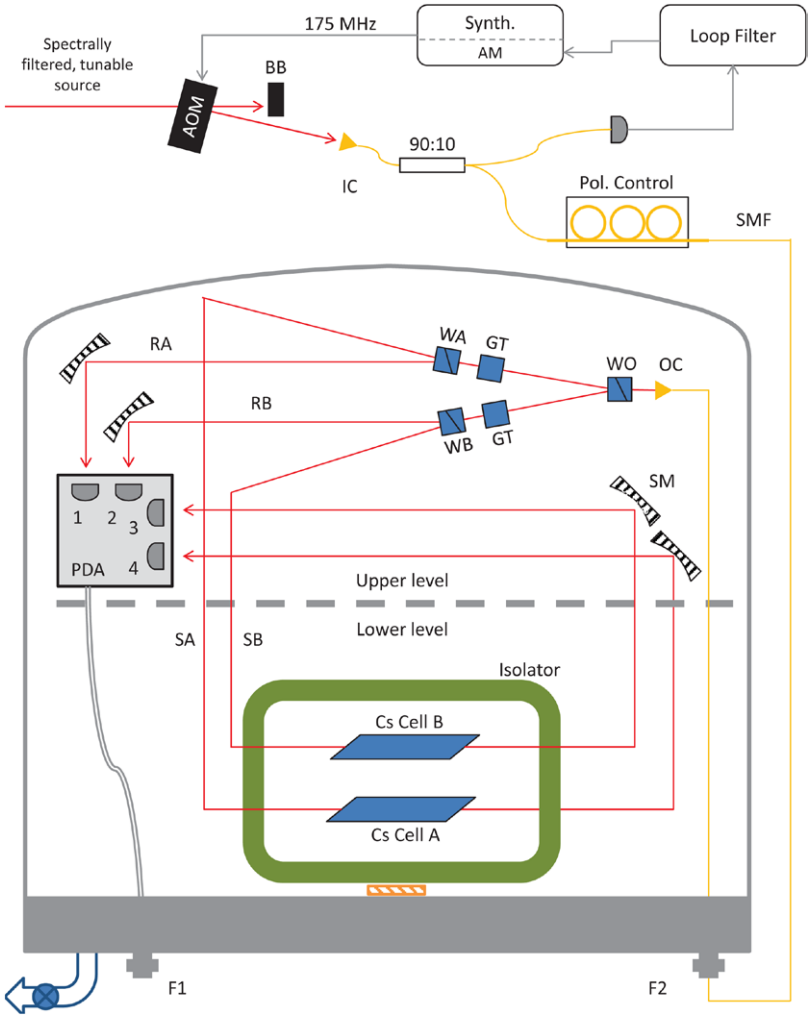

Figure 5. Delivery of signals into the experimental chamber. AOM: acousto-optic modulator, BB: beam block; IC: input coupler; Synth.: frequency synthesiser with remote amplitude control; SMF: single-mode fibre; OC: output coupler; WO, WA, WB: Wollaston polarising beamsplitters; GT: Glan-Taylor prism; RA, RB: reference beams; SA, SB: sample beams; SM: spherical mirror, PDA: photodetector array; F1, F2: vaccum feedthroughs.

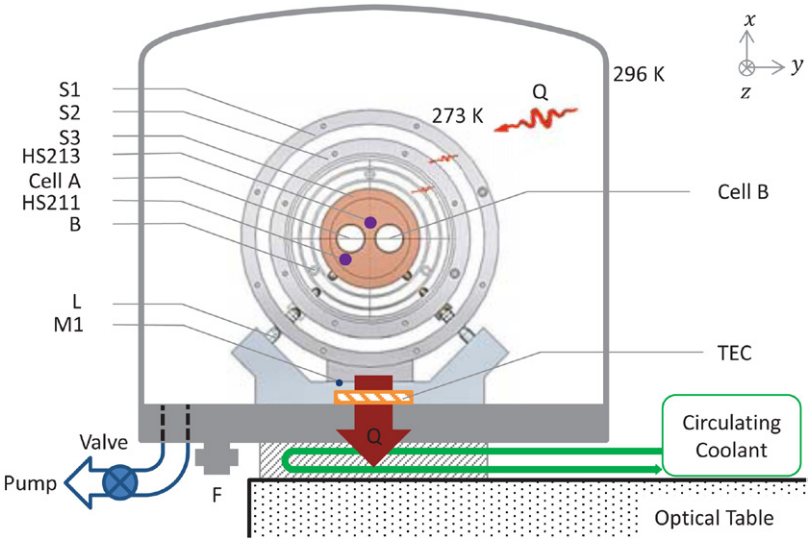

Figure 6. Schematic diagram of heat flows and major components of the environmental shields. Cell A, B: Cs cells; S1, S2, S3: thermal shield layers; HS213, HS211: calibrated CSPRTs; B: dual-layer magnetic shield; L: domed insulating legs, TEC: thermoelectric cooler; M1: monitor thermistor; Q: heat transfer; F: optical and electrical vacuum feedthroughs.

Table 5. Summary of simulated thermal gradients, calculated time constant of each thermal shield and the Cs cells. The cells were assumed to be in good thermal contact with S3 so that its thermal time constant was the same as S3.

Element	ΔT(μK)	$\Delta T/T$(ppm)	τ_{TH}(hr)
S1	200 000	730	43
S2	150	0.5	40
S3	9	0.03	18
Cell	3	0.01	18

Note: The total system can be approximated by a series of cascaded low pass filters with these time constants.

gradient introduced by the asymmetric extraction of heat from this shell. The intermediate shell S2 was thermally isolated from S1 by resting it on highly insulating Teflon legs (L). The gradients sustained on S1 were radiatively coupled to S2, and the goal was to ensure that this inter-shield coupling was slow compared to the internal thermalisation timescale of S2. On this passive shield, a moderate amount of thermal inertia is desirable to filter away residual fast fluctuations propagated through S1. On the other hand, too much inertia will unnecessarily lengthen the time required to reach the desired temperature. The final stage S3 was designed to provide passive stability and to minimise gradients on the cells. For this reason, S3 was fabricated from copper to exploit its higher thermal diffusivity. This was less critical for S1 and S2 so they were made from more readily machinable aluminium. The temperature of S3 was monitored using two calibrated capsule standard platinum resistance thermometers (CSPRTs) labeled HS211 and HS213. Electrical and optical vacuum feedthroughs were used to introduce and extract signals from the chamber. Greater temperature stability than the passive performance of the nested shells could be achieved using closed-loop control of the temperature on S1 by feedback to the TEC to keep a monitor thermistor M1 at constant temperature.

We used a commercial finite elements analysis (FEA) software package [32] to model the transfer of gradients from S1 onto the cells. Heat flow within the shells was simulated using the discretized Heat Equation [33]. Radiative heat transfer was more difficult to describe because the mean free path for radiative transfer is much greater than that for conduction: in these cases the local temperature $T(\mathbf{x}, t)$, at point $\mathbf{x}$ and time t, and its derivatives can depend on $T(\mathbf{x}+\mathbf{a}, t)$ that is spatially separated by a finite distance $|\mathbf{a}|$. The ballistic nature of the photons that transport heat means that geometric details far from $\mathbf{x}$ can affect the radiative flux impinging on a surface at $\mathbf{x}$.

In our highly symmetric geometry, it was straight-forward to simulate the radiative coupling between shells by calculating the emitted flux from the surface of each shell and combining this with their view factors (defined as the fraction of energy emitted at a distant point that is collected at a local point on the shield surface) [34]. Effectively, we captured the physics of radiative heat transport by describing it in terms of a boundary condition, thereby restoring the ability to express $T(\mathbf{x})$ and its derivative in terms of (local) quantities at $\mathbf{x}$. Further details of the mathematical procedures can be found in [35]. The magnitude of the calculated steady-state gradients and the thermal time constants for each shell is summarised in table 5; this facilitates a lumped-mass analysis [36].

4.2. Performance benchmarking

Two capsule-type standard platinum resistance thermometers (CSPRTs) manufactured by Hart Scientific (Model 5686, SN:HS211 and HS213) were calibrated to the ITS-90 temperature scale [37] by the National Institute of Standards and Technology (NIST), MD, for the temperature range between $T = 84$ K to $T = 430$ K. The total temperature uncertainty in the calibration was less than 0.1 ppm for the range $270\,\mathrm{K} < T < 300\,\mathrm{K}$ and so the CSPRTs were used as a transfer temperature standard in this experiment. We note, however, that the ITS-90 scale itself has been shown to underestimate true thermodynamic temperature by 4 ± 1.4 ppm at 296 K [38]. By definition, the ITS-90 scale agrees with true thermodynamic temperature at the triple-point of water, and this motivated us to also perform the experiments near to this point.

A high-precision Hewlett-Packard Model 3458A (SN:2823A-03101) digital multimeter (DMM), which was calibrated by the National Measurement Institute of Australia, provided an inferred temperature measurement from this device with 1.2 ppm total uncertainty. In principle, there is no impediment to matching the CSPRT calibration uncertainty (0.1 ppm) using commercially available resistance bridges (such as the ASL Model F900 which offers 0.02 ppm accuracy when correctly calibrated) [39].

Conceptually, the simplest method to measure thermal gradients is to synchronously measure the local temperature at two locations on S3. To this end, one of the CSRPTs (HS211) was located at the centre of S3 and functioned as the primary monitor of the sample cell temperature. An auxiliary sensor (HS213) was located at the largest practical separation from HS211 to search for thermal gradients. The locations of the SPRTs are indicated on figure 7. Synchronous measurements were not possible because only one high-precision DMM was available. Instead, we rely on the high thermal inertia of the isolator to make an inferred measurement of the worst-case gradient. Figure 8 shows the temperature of S3 as measured at the locations of HS211 and HS213 just before the isolator reaches equilibrium after an earlier step change in set point. The temperature at the centre of S3 is recorded for ∼15 hr (red trace) and extrapolated for the next 0.5 hr. Meanwhile, the DMM input is switched to measure the temperature at the edge of S3. Comparison of the extrapolated temperature at the centre of S3 to that of the edge places an upper bound on the temperature difference across the cell (about 75 mm) to be less than 0.2 ± 0.2 mK, corresponding to a fractional temperature error of 0.7 ± 0.7 ppm.

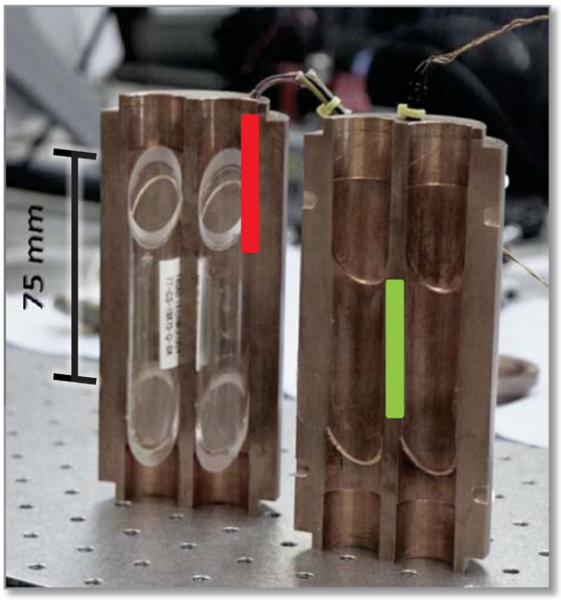

Figure 7. Photograph of S3, opened to show the locations of the HS213 (green) and HS211 (red) CSPRTs and the Brewster-angled Cs sample cells.

The Allan deviation of the temperature at the centre of S3 (figure 9) shows that the measurements are limited by the DMM input noise for short time up to ∼60 s, before a drift behaviour becomes dominant at longer timescales. The large thermal inertia of S3 prevents temperature fluctuations on a timescale less than 18 hr (table 5), permitting us to assume that the measured drift rate at large integration times can be extrapolated to infer the stability at shorter times. The two timescales of interest were (I) time between successive spectroscopic measurements (<1 s), and (II) time to acquire a single spectrum (∼60 s). We measured variations in temperature no larger than 0.2 mK (0.7 ppm) which is at the same level as the upper bound set by thermal gradients. Other work has found that the typical SPRT response drifts with time at the level of 0.8 mK yr^{-1} [40]. We do not explore this source of systematic error, but in principle, a recalibration immediately prior to a set of experiments may be used to correct for any calibration drift.

5. Magnetic isolation

In the Cs D1 transitions that we probed ($F = 3$ to $F' = 3, 4$), the worst case shift of an unresolvable Zeeman transition (1.41 MHz G^{-1}) was realised by the so-called π transitions for $|m| = 3$ [41]. We built a magnetic shield to suppress the maximum frequency shift to 1/10 of the uncertainty in the atomic Lorentzian linewidth (i.e. 6 kHz/10 = 600 Hz). i.e. the maximum tolerable external field passing through the cells is 0.5 mG. Since the measured laboratory field strength at the location of the experiment was less than 0.5 G, we required an attenuation factor in the magnetic field of ∼1000. A schematic view of a dual-layer magnetic shield that was designed to fit in between thermal shields S2 and S3 is shown figure 10. The body comprises of two nested coaxial cylinders with outer diameters 100 mm and 84 mm of length 220 mm and 210 mm, respectively, and thickness 1.5 mm. The two cylinders were rigidly attached to each other with stainless steel (low permeability) spacers. The optical probe beams were introduced through 20 mm apertures in the end caps, which friction-fit onto the cylindrical bodies with overlapping collars. Smaller apertures were available for the CSPRT signals to pass through the shields. Every aperture was guarded by a tubular extension on the lids of the inner shield.

A finite element analysis approach was again used to simulate the worst-case performance of the magnetic shield in which the external field was oriented parallel to the cylinder axis. For convenience, the shields were cylindrically symmetric so that the cells and apertures were centred on the cylindrical axis. We expect this to be an inconsequential approximation because we have only laterally translated those features in a uniform field. In the region occupied by the Cs cells the predicted attenuation factor exceeds 31 600, (figure 11) which is an order-of-magnitude greater than required.

The anticipated reduction factor was challenging to confirm experimentally: a background field of 50 μT (0.5 G) would be reduced to 1.6 nT, which was below the DC resolution of our Hall-effect gaussmeter (Sypris Model 6010) that was around 10 μT. Instead, we placed the magnetic shield assembly coaxially inside a solenoid consisting of 430 turns. A low frequency (7 Hz) oscillating current was passed through the solenoid to generate a uniform external magnetic field aligned to the cylindrical axis of the shields. This is the worst-case orientation for leakage of magnetic flux into the shielded volume via the laser access apertures. The largest applied field before the shields became magnetically saturated was approximately 1 mT. Under these conditions, the field inside the chamber could be reduced to approx. 600 nT, which was the limit set by a synchronous measurement with a lockin amplifier. We were thus able to verify that the worst-case shielding factor was around 1600, which exceeds the minimum design specifications.

6. Alkali absorption spectroscopy

6.1. Absorption spectroscopy on rubidium D2 transition

In figure 13 we show data (red curve) obtained in our first generation spectrometer on the Rb-D2 transition for a natural mixture of Rb-85 and Rb-87 along with the theoretically expected lineshape. The residuals are at a level of 10^{-3} and show a number of distinct features: the broadband white noise is associated with insufficient common-mode rejection of amplitude noise and a higher measurement noise floor relative to our second-generation apparatus; the asymmetric features on the Rb absorption features are associated with residual frequency noise on the probe laser. Each of the spectral features seen on this curve are actually multiple resonances that are not resolved (as seen on figure 2). This made the fitting procedure unreliable as there were a number of correlated values in the fitting algorithms. We note also that this experiment made use of synchronous detection with

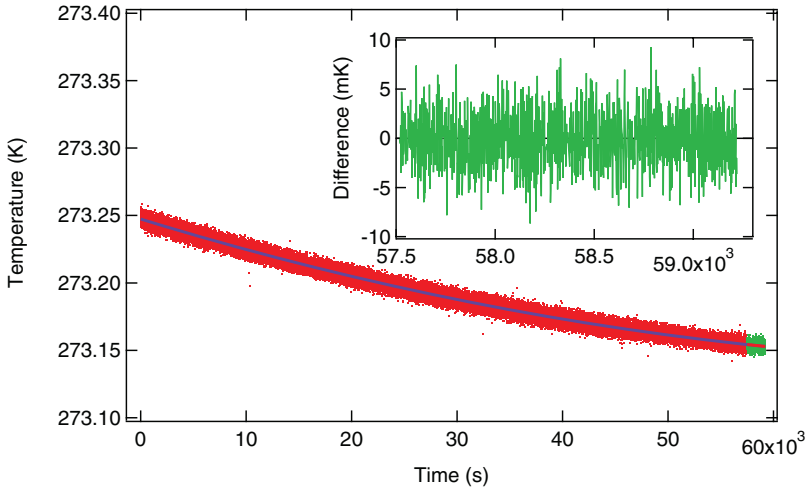

Figure 8. The temperature gradient across S3 was inferred by recording the temperature at the centre (red dots) and extrapolated (red line). The measured temperature at the edge of S3 (green dots) were compared to the extrapolated values and the difference is shown in the inset.

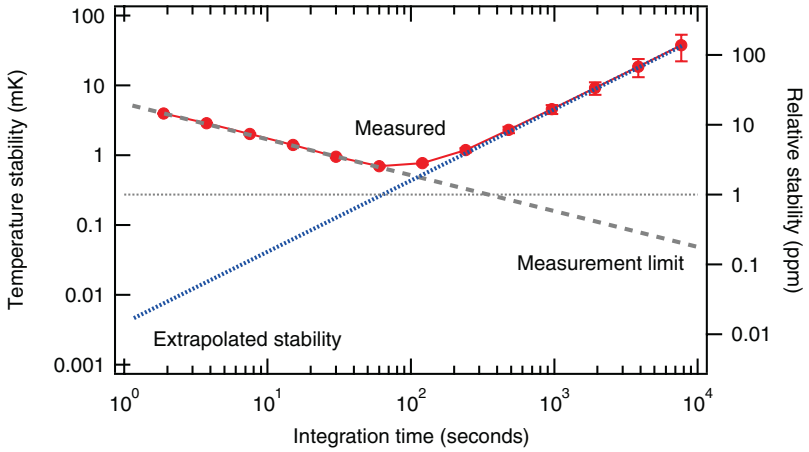

Figure 9. Temperature stability measured at the centre of S3 (red circles) was measurement limited by the HP3458A DMM at low integration times (grey dashed). The expected stability was extrapolated (blue dashed) by assuming a constant drift rate.

lockin amplifiers to read out the detectors to avoid long-term drifts and background light fluctuations. We found that the nonlinearity and dynamic range limits of the lockin amplifier input stages prevented a high precision experiment. We thus dispensed with this approach and turned to a new improved spectrometer (described below), along with replacing Rb with Cs, to resolve these issues.

6.2. Absorption spectroscopy on cesium D1 transition

We measure the transmission through a Cs vapour cell as the probe laser frequency scans across the D1 transition ($6S_{1/2}$–$6P_{1/2}$), shown in figure 14. The expected lineshape is that of two Voigt profiles separated by the upper state hyperfine splitting and the experimental data is superficially in accord with this (see lowest panel of figure 14). By fitting a line-shape of the form of equation (4) to the measured transmission data we can extract the physical parameters of the resonances. One of these extracted parameters, Γ_D^{fit}, gives us the Doppler component, from which we infer the temperature. We can then characterise the systematic errors in Γ_D^{fit} by comparing to the independent PRT temperature measurements, Γ_D^{PRT}.

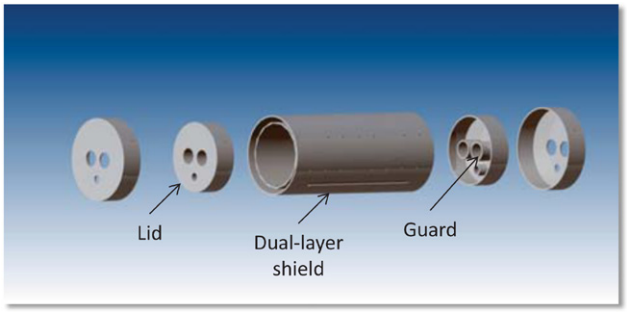

Figure 10. Rendered drawing of the dual-layer magnetic shield assembly showing the guards around the apertures required for laser and electrical access.

We show the residuals from fitting the average of 200 scans at the highest explored powers (600 nW) on the second lowest panel of figure 14) and we see immediately deviations from the expected Voigt profile at the 200 ppm level. Some of these deviations are technical in origin e.g. instrumental broadening due to the lineshape of the probing laser, residual spontaneous emission from the probe laser and photodetector linearity; however, there is an important fundamental

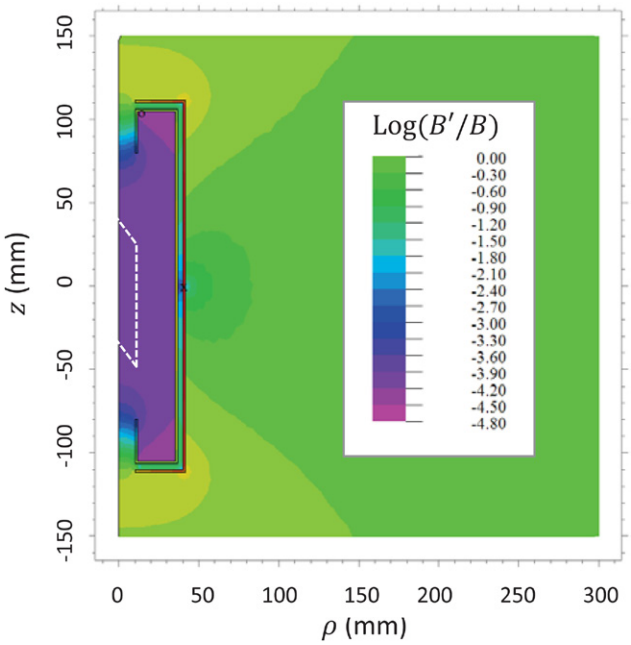

Figure 11. Contour plot of the attenuation of an externally applied field parallel with the z axis. Cylindrical symmetry has been assumed for convenience. The location of the Cs cell is outlined (dashed white) where the attenuation exceeds 31600.

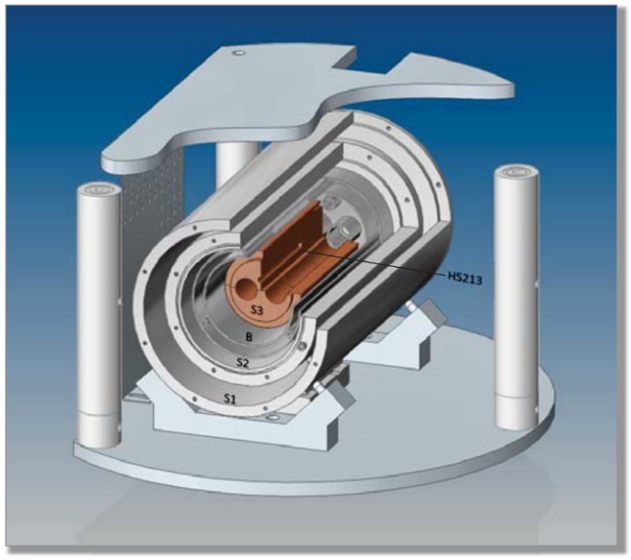

Figure 12. Exploded view of the vacuum system along with thermal and magnetic shielding.

effect resulting from frequency-dependent optical pumping, which perturbs the natural line-shape so that it is no longer a Lorentzian function. When this modified natural lineshape is convolved with the Doppler broadening it gives rise to a lineshape that is not a Voigt function; it is exactly this effect that gives rise to the strong residuals seen at frequencies in the environs of the atomic absorption features. We note that there are also other modulations (at the 30 ppm level) in the residuals at frequencies that are far detuned (>2 GHz) from the atomic resonances. These are associated with unavoidable optical etalons in the measurement system as we will show below. All of these effects, whether technical or fundamental,

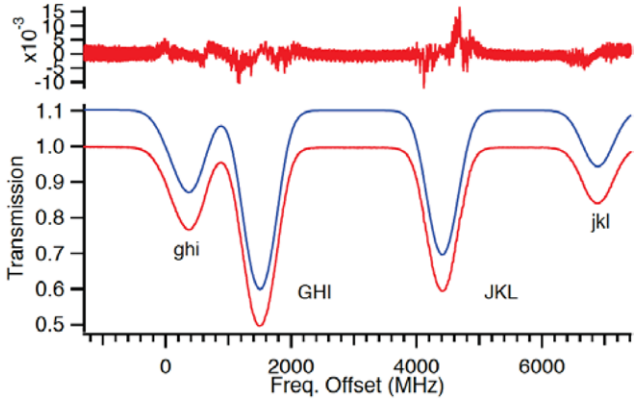

Figure 13. A broad-band transmission spectrum of Rubidium vapour—red curve is experimental data while blue curve is theoretically expected shape (intentionally moved upwards for clarity). The residuals are seen at the level of 10^{-3}.

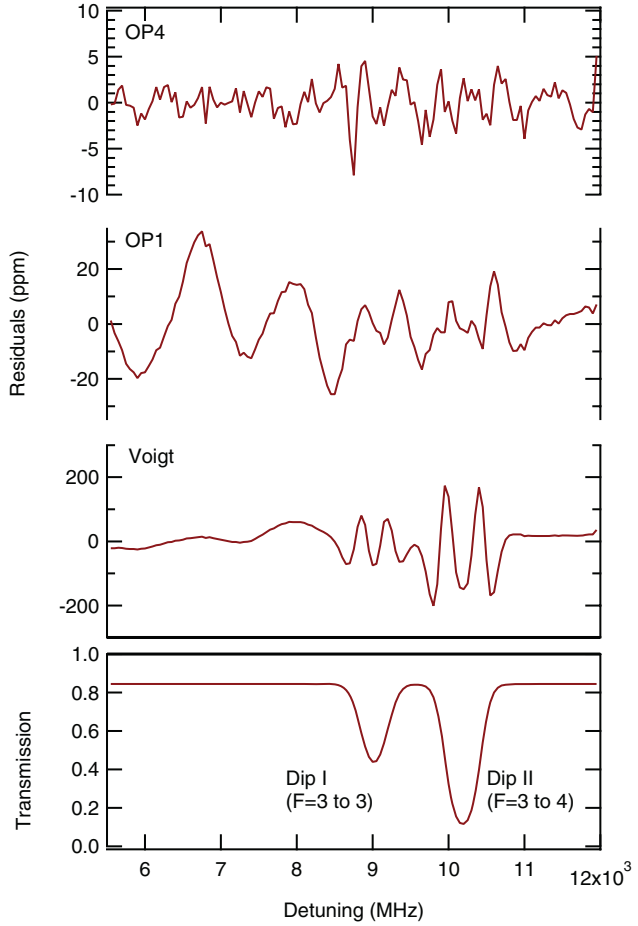

Figure 14. Typical Cs spectra (averaged over 200 sweeps) taken at 296 K (lowest panel), and the residuals to least-squares fitting using various models; Voigt: linear absorption model with purely Voigt profiles, OP1: optically-pumped model with a no etalons terms in the model, OP4: optically-pumped model with four etalons included.

cause systematic perturbations to the line shape, and must be taken into account to model the lineshape accurately (and to obtain a high-quality estimate of Γ_D^{fit}).

We have developed a model that allows us to include the optical pumping effects on the atomic lineshape [25, 31].

When we include this more more complex line shape in the fitting model we obtain the residuals labelled 'OP1' on figure 14; the residuals are now much smaller and we have increased the residual scale by 10 fold to see them more clearly. Once we have effectively removed the spectral features associated with optical pumping then the dominant features are associated with the unwanted etalons that contribute across the entire scan (including at large detunings from the atomic resonances). These etalons, which arise from stray reflections, either specular and diffuse, are a serious issue in any laser experiment aimed at measuring ratios of power levels. This is of course a necessity in an experiment aimed at measuring transmission.

The dominant etalon has an amplitude around 30 ppm and is believed to be formed through interference of the main beam with scattered light rather than any specular reflection (see below for more details on this). Although obviously technical in nature, reducing the size of etalons beyond this already fantastically low level is a significant experimental challenge. A thorough discussion of the sources and diagnostic procedures used to minimise the effect of etalons is included in the Supplemental Information section of [31] and below. To handle these residual etalons in the data reduction, we include etalons in our transmission model so that the total transmission is given by $\mathcal{T} = \mathcal{T}_{\rm at}\mathcal{T}_{\rm et}$, where $\mathcal{T}_{\rm et}$ includes n etalons and a slowly-varying quadratic background:

$$\mathcal{T}_{\rm et}(f) = \frac{\alpha + \beta f + \gamma f^2}{\prod_{j=1}^{n}(1 + a_j \sin(2\pi f/f_j + \phi_j))}. \quad (5)$$

If we include four etalons ($n = 4$) together with the more complex optical pumping model we obtain a residual transmission measurement labelled 'OP4' on figure 14). The relative noise in the measurement of the atomic transmission is now just 2 ppm in a 1s measurement which is within a factor of 2 of the photon shot-noise in the experiment. It is only necessary to include 3 etalon terms for the retrieved value for Boltzmann's constant to be consistent with the known value—adding additional etalon terms beyond this does not change the retrieved value within the precision of this experiment (we explain this in more detail below). We have chosen to shown the $n = 4$ residuals on figure 14 because the transmission noise is a little better.

We have made a total of 1000 scans of the Cs transmission at 5 different incident powers stretching from 20 nW to 600 nW incident power in a 0.5 cm probe beam. We used these exceedingly small powers to limit the effects of optical pumping.

As an example of the type of data extractable from these fits we display on figure 15 the results of the frequency splitting between the two absorption resonances, which corresponds to the upper state ($6P_{1/2}$) hyperfine frequency splitting. The frequency splitting shows no evidence of power dependence as would be expected. The statistical deviations for $k_{\rm B}$ reported by the fit at the various powers shown on figure 15 are 36 kHz; 11 kHz; 7 kHz; 4 kHz and 3 kHz. The sample deviation for the same sets of data are 32 kHz; 10 kHz; 8 kHz; 5 kHz and 4 kHz

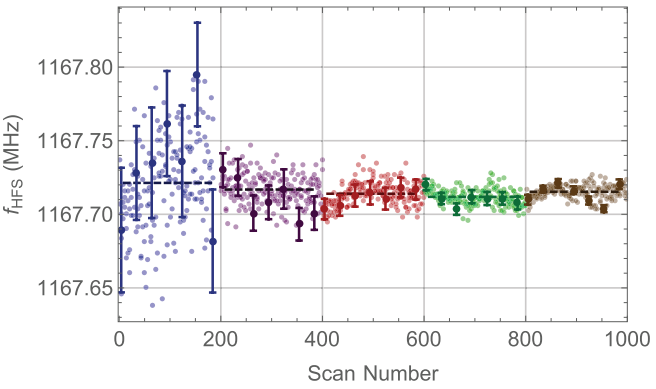

Figure 15. The extracted frequency splitting between the two absorption resonances over 1000 separate scans. The different colours ranges correspond to different incident powers: 28 nW (blue); 97 nW (purple); 186 nW (red); 332 nW (green); 858 nW (brown). For seven points in each power range we have shown the uncertainty on the extracted value as reported by our fitting algorithm. We see that for the lower powers this is closely consistent with the scatter of the multiple points. At the higher powers we see evidence of some systematic problems at the level of ~5 kHz.

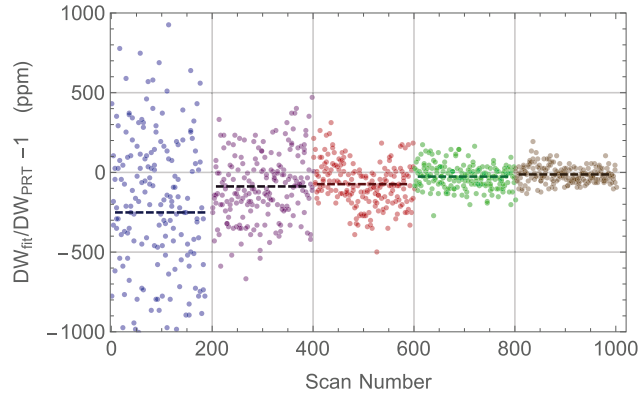

Figure 16. The extracted Doppler component of the absorption resonances over 1000 separate scans. The different colours ranges correspond to different incident powers: 28 nW (blue); 97 nW (purple); 186 nW (red); 332 nW (green); 858 nW (brown). The dashed line shows the mean value at each incident power level which shows no significant statistically significant shift except at the lower incident power (where we think the poor signal to noise ratio of the data led to problems in correctly identifying background etalons.

showing an absence of any unexplained systematic effects for all but the most high resolution measurements. The slight difference in the statistical and sample deviations for the highest powers is perhaps arising from some small residual slope coming from background etalons, however the corresponding influence of this on the Doppler width is below the 1 kHz level, which would not affect the extracted value for the Boltzmann constant at our current level of precision.

On figure 16 we show the difference between the extracted Doppler width component of the absorption features and that predicted from the sample temperature which was independently measured with a calibrated platinum thermometer. We see that the implementation of our optical pumping model suppresses any statistically significant variations in that extracted parameter across this incident power range (except for the

Table 6. Experimental error budget for the determination of the Boltzmann constant at 296 K with a 2 h measurement campaign.

Source	Current $u_r(k_B)$ (ppm)	Upgrade $u_r(k_B)$ (ppm)
Statistical	5.8	0.1
Lorentz Width (Γ_L^{fit})	65	1.5
Laser Gaussian noise	16	0.001
Optical pumping	15	1
Etalons (misidentification)	15	0
Etalons (unresolved)	3	0.1
Spontaneous Emission	3.6	0.2
Temperature	1.9	0.2
Temp. Gradient	1.2	0.2
PD Linearity	1	1
Zeeman Splitting	<0.1	<0.1
Atomic recoil	<0.1	<0.1
Total (fit Γ_L)	71	2.0

Note: Optical pumping shifts are reported at $I/I_{sat} = 3 \times 10^{-3}$ assuming second-order corrections to the Voigt profile. u_r denotes fractional uncertainty. Second column shows an upgraded spectrometer with a narrower linewidth probe laser that can have its output power switched between high and low levels (see text for details) and over a month-long measurement.

lowest powers where we suspect residual etalons are influencing the values recovered by the fit because of the limited signal to noise ratio). On the other hand, if we use a simple Voigt profile to fit the data then we see a physically unrealistic 800 ppm shift in the extracted value for k_B at the highest powers when compared to the lowest powers. This increases our confidence in our optical pumping model because it not only fits the measured spectral lineshape better (evidenced by the reduction in the residuals) but also eliminates the physically unrealistic dependence of k_B on the incident intensity of the probe.

The statistical and sample deviations for k_B are consistent across the various powers (i.e. we do not see any unexplained systematic effects) and correspond to (first number is statistical error generated by the fit while bracketed number is the sample deviation): 495(692) ppm; 208(225) ppm; 139(135) ppm; 76(78) ppm and 53(53) ppm for the same incident power ranges shown on figure 16.

By taking a weighting average of all the scans we find $k_B = 1.380\,545(98) \times 10^{-23}$ J K^{-1} where the 71 ppm uncertainty is calculated in table 6 with the dominant uncertainty associated with the Γ_L. The result for k_B is consistent with the recommended CODATA value of $1.380\,6488(13) \times 10^{-23}$ J K^{-1}.

7. Current and ultimate limits to atomic Boltzmann thermometry

In this section we will consider the limits to Doppler broadened spectroscopy set by technical and fundamental processes. These processes impose limits on both the precision and accuracy of the measurement. The current and potential limits are listed in table 6 and will be described in more detail below. We term some of these as technical issues (e.g. etalons) although they are essentially unavoidable in any real system and hence need to be considered just as seriously as fundamental limits. In the interests of brevity we have ignored a number of potential contributors below where we have previously shown them to contribute at the part-per-billion level or below: these include photo-association, recoil, quantum interference and quantum limits which are addressed elsewhere [31, 42].

Below we describe the relative uncertainty of a parameter x as $u_r(x)$—this represents the standard deviation of the extracted parameter divided by its value.

7.1. Technical limits

Detection noise and Linearity: Reverse-biased silicon photodiodes were used to measure the optical power in each of the reference and sample beams. The photosignals were measured by digital multimeters (DMMs) recording the voltage generated by the photocurrent passing through a load resistor. The contributions from shot noise, Johnson noise from the load resistor and the DMM input voltage noise were characterised [31]. At the highest powers used in the experiments (∼600 nW), shot noise contributed 50% of the total measurement noise although in all the explored power ranges it was greater than 10% of the total noise.

The linearity of the photodetection chain was measured in two ways. In the first approach, the ratio of the transmitted and incident detector signals were measured over the full range of input power variation. This provided a reliable calibration curve for the transmission measurement which allowed deviations from linearity at high powers to be corrected. For low powers, this method did not provide sufficient signal-to-noise for correction of small offset voltages. To determine these offsets, we made use of the actual atomic absorption in the two available transitions to provide an *in situ* calibration. The Clebsch–Gordon coefficients for these transitions require that the linear (i.e. low-power) ratio of absorption coefficients should be 1:3 [41]. Using the extracted absorption depths of the two transitions we determined the offset levels required to make this ratio agree with the 1:3 condition as we the incident power on the cell is reduced towards zero. We thus have a guaranteed and independent way to determine any nonlinearity in the detector response across the full power range. The technique allowed us to measure offsets corresponding to <8 pW thus ensuring linearity over a dynamic range of 10^5 : 1. A more complete description of this method is found in [43].

Probe Laser Purity: In a separate experiment the slave laser light was tuned to the centre of the atomic resonance and passed through a heated Cs cell for which the optical depth was 30 i.e. $\mathcal{T}_{at} \approx e^{-30} \approx 10^{-12}$. A measurement of the residual transmission demonstrated that the laser output contained 1.6% broadband emission at its operating point. It was thus necessary to filter the laser output spectrum because broadband emissions contaminate the measurement by creating a transmission offset associated with the non-resonant photons.

The laser light was filtered using the fundamental mode of a scannable Fabry–Perot cavity ($\mathcal{F} = 305$ and free-spectral range 17.6 GHz). Low level feedthrough of light from higher-order spatial modes of the optical cavity (∼20% of the main mode) was heavily attenuated by coupling the beam into a single-mode optical fibre prior to introduction into the evacuated chamber. The combined cavity and single-mode filtering provides a reduction in the spontaneous emission by a factor

$2\mathcal{F}/\pi \approx 195$. We numerically modeled the effect of the residual spontaneous emission (0.008%) on the measured Cs spectrum and found that the perturbation to the fitted atomic Doppler width ν_D was only 1.8 ppm (equivalent to 3.6 ppm in $u_r(k_B)$). This systematic error can easily be reduced to the 0.1 ppm level with a readily achievable cavity finesse of over 10 000. One should note that it will be necessary to reduce the fast laser linewidth in order to avoid intensity noise from being produced in transmission through this high finesse cavity (for example, by using the laser described below).

The observed transmission feature is a convolution of the probe laser spectral characteristics and the underlying atomic absorption. In order to obtain a good representation of the underlying atomic characteristics it is necessary to deconvolve the effects of the probe laser spectrum. It has been shown that diode laser spectral lineshapes can be expressed to a high degree of accuracy as a Voigt function [44]. Let us define the Lorentzian half-width-at-half maximum of the laser as Γ_L^{las} and the Gaussian e^{-1}-halfwidth as Γ_D^{las}. When this approximation is convolved with the atomic absorption profile (which, to first order, is also a Voigt function with parameters Γ_L and Γ_D), we obtain an observed atomic lineshape that is also a Voigt function with Lorentzian and Gaussian components given by $\Gamma_L = \Gamma_L^{at} + \Gamma_L^{las}$ and $\Gamma_D = \sqrt{(\Gamma_D^{at})^2 + (\Gamma_D^{las})^2}$, respectively.

The power spectral density (PSD), $S_\nu(f)$, of the probe laser frequency noise was directly measured by using the frequency-dependent absorption on the side of an atomic transition. Two regions of the PSD can then be identified [44]; at higher Fourier frequencies we see a dominant white frequency process, while at lower frequencies we observe steeper noise which in our case has a dominantly flicker ($1/f$) character. White noise components (with amplitude S_0) produce a Lorentzian lineshape with HWHM given by $\pi S_0/2$. The integrated probe laser noise at lower frequencies give rise to a Gaussian lineshape which has an estimated e-fold width of $\Gamma_D^{las} = 0.88(39)$ MHz when integrated over the 40 ms observation time for each frequency point in the scan. The Gaussian component sums in quadrature with Γ_D, and so contributes 8 ppm to $u_r(\Gamma_D^{fit})$ [16 ppm to $u_r(k_B)$].

We estimated the Γ_L^{las} independently by making a beat-note between two nominally identical probe lasers and measuring the noise at high offset frequencies—this gave a linewidth of $\Gamma_L^{las} = 44(18)$ kHz. We can then fix the value of Γ_L in the fit to the sum of laser noise and the atomic linewidth i.e. $\Gamma_L^{at} + \Gamma_L^{las} = 2.287(6)\,\text{MHz} + 44(18)\,\text{kHz} = 2.331(18)$ MHz [41]. A better alternative is to treat Γ_L as a free fitting parameter (together with Γ_D) and then we find that we can extract a value for Γ_L^{las} of 2.327(7) MHz which is consistent with the earlier value but nearly twice as precise. Nonetheless, the uncertainty of this better approach still dominates our estimate of Γ_D^{fit} because of the correlated nature of the Lorentzian and Gaussian components in the Voigt.

It is worth considering the potential noise-floor of this approach if we were to use a higher quality tunable laser. A reasonably high-quality fibre-stabilised laser shows a Lorentzian linewidth of around 0.1 Hz and a Gaussian linewidth of ~2 Hz [45]. The use of such a laser would reduce the uncertainty contribution to k_B from the spectral purity of the laser to below 1 ppb. As we show below, this approach could not only provide an excellent measurement of Γ_D but simultaneously generate a value for the atomic lifetime, Γ_L^{at} that is 10 fold better than the current state of the art.

Conventional thermometry: We were able to show no temperature gradients above the 1.2 ppm level, which was limited by our resistance meter (HP Model 3458A). The ITS-90 temperature scale, T_{90}, itself is known to deviate from the true thermodynamic temperature, T by $T - T_{90} = 3.2 \pm 0.4$ mK at $T_{90} = 296$ K [38]. When the uncertainty in this correction is included, the total uncertainty due to conventional thermometry is 1.9 ppm. There are readily deployable techniques to upgrade this performance by a factor of 10.

Measurement of residual B fields: As outlined earlier we are able to demonstrates that the minimum magnetic shielding factor meant that the worst-case Zeeman splitting remained below $1/10^{th}$ of the Lorentzian linewidth. A simulation of the Zeeman splitting on the recovered Doppler width shows that this shift is less than 0.1 ppm. By increasing the magnetic shielding it is clear that this effect could be reduced arbitrarily.

Etalons: Appropriate treatment of these etalons are of particular concern for DBT since they have the potential to introduce systematic errors in Γ_D^{fit}.

The largest etalons have an amplitude of ~3–40 ppm in the transmission ratio $\mathcal{T}$ over the ~6 GHz spectral band we sampled. These etalons appear in spite of some substantial efforts taken in the experiment to eliminate them; we will demonstrate that very modest amounts of stray light will lead to these levels of baseline imperfections and are difficult to completely eliminate under these high signal-to-noise detection scenarios.

In order to produce an etalon of this magnitude it is only necessary to interfere two fields that have an amplitude ratio of 40 ppm. This corresponds to a power ratio of the two fields of 10^{-9} emphasising how challenging this problem will be in any real experiment. Of course, even the best anti-reflection (AR) coatings have reflectivity thousands of times larger than this. We thus cannot rely on AR coatings but instead make use of a number of techniques to reduce reflection: ensure that there are no parallel surfaces in the experiment, use of carefully adjusted Brewster angle surfaces, and use of low scattering surfaces. Nonetheless, a surface roughness of just a few nanometres on an optical surface is sufficient to scatter a few thousand ppm of the main beam, and for realistic geometries (30 cm spacing between scatterer and collection) the overlap with the main beam will produce 40 ppm of modulation [31]. This cannot be overcome by a slight angling of the surfaces because the scattering occurs into a wide cone. The simplest solution to this problem is to use well-spaced optics to prevent the effect of the scattering processes. Unfortunately, in the tight confinement of the vacuum chamber, this scattering problem was exacerbated, which required inclusion of etalons in our fit model. It is important to note that it is only because of our extremely low transmission noise that we can see these effects at all; but whether visible or not they still cause systematic offsets in obtaining a value for Boltzmann's constant.

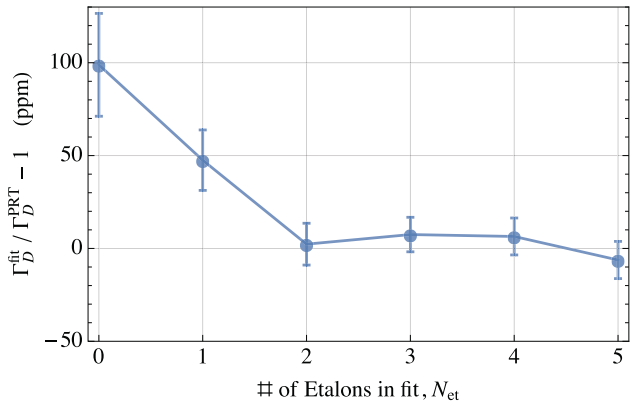

Figure 17. Dependence of Γ_D^{fit} on number of fitted etalons that are included in the fit model. Error bars are standard errors reported by the fitting. After 2 etalons are included in the fit, additional etalons do not significantly change the value of Γ_D^{fit}.

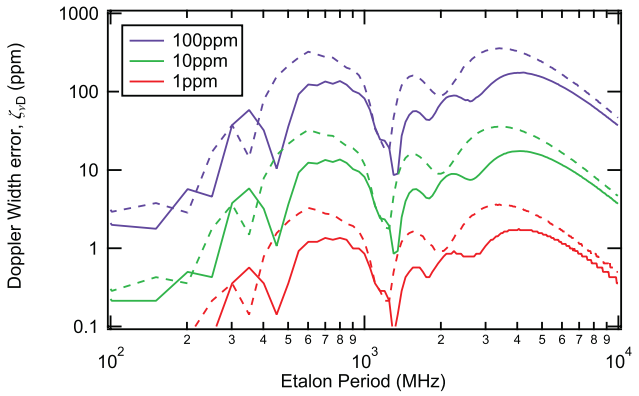

Figure 18. Estimated maximum relative error in the Doppler width caused by a background etalon unaccounted in the model function, plotted as a function of the etalon period. The less absorptive hyperfine transition from $F = 3$ to 3 (dashed traces) is more sensitive to etalons than the deeper $F = 3$–4 transition. Both sensitivities scale linearly with etalon amplitude.

As more etalons are included in the model we are able to more effectively suppress the modulations seen in the transmission at frequencies away from the atomic resonances (see supplementary information in [31] for extensive details). The fitting algorithm can take some substantial time to find the parameters for the model and thus to limit the duration of this fitting process we typically find the etalon parameters by fitting to the average of 200 scans. The values for the etalon parameters are then held fixed for the individual fits to each scan where the other parameters are left free. Figure 17 shows the effect on the extracted value for Γ_D^{fit} as we adjust the number of etalons included in the fit. We see that the extracted value for Boltzmann's constant stably approaches a fixed value, which is also close to the expected value, as we add two etalons and then holds a steady value within the uncertainty of the estimate after that. This gives us confidence that the process of adding etalons is robust, that it is the correct thing to do and, that even if too many etalons are included it will not change the result at a significant level.

In this next section, we consider in greater detail the potential effect of etalons on the final result. They can enter into the uncertainty budget in two ways: first, etalon parameters may be misidentified during fitting, and so contaminate Γ_D^{fit}. Second, unresolved etalons masked by measurement noise may introduce residual systematic shifts to Γ_D. We show both these effects explicitly in Table VI.

7.1.1. Perturbation to estimated doppler width. We briefly modeled the sensitivity of the fitted Doppler width to the presence of etalon backgrounds using simulated spectra. Assuming small amplitude residual etalons, we can implement them into our fit function as follows:

$$\mathcal{T}_{fit}(\nu) = \mathcal{B}(\nu)\mathcal{T}_{at}(\nu), \tag{6}$$

$$\mathcal{B}(\nu) = \sum_{k=0}^{p} A_k \nu^k + \sum_{j=0}^{q} B_j \cos\left(2\pi \frac{\nu}{\Pi_j} + \Phi_j\right) \tag{7}$$

where $\mathcal{T}_{at}$ is the ideal atomic transmission function. We allowed a polynomial variation in the background, which depended on the coefficients A_k, in addition to the etalon terms which have amplitude B_j, period Π_j and phase Φ_j.

The perturbation to the estimated Doppler width caused by a residual etalon will depend not only on the etalon amplitude, but also its period and phase (i.e. Π_j and Φ_j). For example, an etalon whose period is much broader than the Doppler width of the Cs resonances would likely contribute an offset in the background level that is effectively captured by the fit parameter A_0. Conversely, a narrow etalon ($\Pi \ll \nu_D$) is averaged-over in the fitting process. For etalon periods that are commensurate with twice the Doppler width (~400 MHz) and the upper state hyperfine splitting (~1 GHz), the perturbation to the fitted Doppler width becomes strongly dependent on the phase Φ of the etalon with respect to the location of the atomic resonance. The potential maximum bias to the fitted Doppler width was determined by simulating etalon-contaminated Cs spectra of known Doppler width that were subsequently fit using the model equation (6) with $p = 2$ and $q = 0$ (i.e. deliberately excluding etalon terms from the fit).

Figure 18 shows the Doppler width error as a function etalon period. The maximum error over all possible etalon phases is reported for each etalon period. The Doppler width error ζ_{ν_D} was defined to be

$$\zeta_{\nu_D} = \nu_{D,fit}/\nu_D - 1 \tag{8}$$

where $\nu_{D,fit}$ is the fitted value for the expected Doppler width ν_D. As expected, there is enhanced bias in the the fitted Doppler width to etalons of periodicity similar to the Doppler width and to the hyperfine separation. Three etalon amplitudes were investigated (100, 10 and 1 ppm) for the optical depths observed for the two Cs resonances at room temperature ($\alpha_{33} \approx 0.66$, $\alpha_{34} \approx 2$). This demonstrated that the sensitivity of the Doppler width error is approximately proportional to the etalon amplitude and optical depth. This is readily understood; the larger the height of the Cs resonance relative to the etalon amplitude, the smaller the disturbance the etalon on the atomic fit parameters.

We have fitted all 1000 transmission scans with a model including three etalons and the optical pumping model. The

weighted average of the standard errors is 8 ppm in $u_r(\Gamma_D^{fit})$, which is a convolution of the statistical noise discussed above and uncertainties arising from possible misidentification of the etalon parameters during fits. When the statistical errors in Γ_D^{fit} are de-convolved from the standard errors, we find the systematic error due to free etalon parameters in the fitting routine to be $(8^2 - 2.9^2)^{1/2} = 7.5$ ppm in $u_r(\Gamma_D^{fit})$, (15 ppm in $u_r(k_B)$).

We bound the error arising from unresolved etalons in two ways: by synthesising data with false etalons, and by computing the fractional shift in Γ_D after the addition of the j^{th} etalon in the fit function. Both approaches give worst-case shifts in Γ_D^{fit} that are comparable to the amplitude of the etalon. The largest unresolved etalons are smaller than the measurement noise floor, so these cannot contribute more than 3 ppm to $u_r(k_B)$.

An upgraded experiment would use a probe laser whose output power could be modified as it scans across the transmission spectrum. We would increase the optical input power when detuned from atomic resonance, and would also scan a larger range of frequency to measure lower frequency etalons. A simulation shows that the application of a narrow-linewidth 1 mW laser tuned over 20 GHz would reduce the effect of etalons of Γ_D^{fit} below 0.1 ppm. This is achieved through both etalon terms being substantially lessened: at the moment we need to average over a number of scans to obtain sufficiently good signal to noise ratio to properly identify small etalons—during this time the etalon parameters slowly drift and this leads to a failure to find the exact correct etalons in any particular scan (this is the misidentified etalon problem). A higher power scan will allow direct identification in a single scan which eliminates this term from our error budget. Furthermore, we can detect an even lower amplitude etalon through the higher power and larger scan range. We estimate that the maximum amplitude etalon is around 0.1 ppm amplitude which, through figure 18, we can see could only induce a maximum systematic error in Γ_D^{fit} of a similar magnitude.

Cell Contamination Another experimental imperfection that perturbed the Cs linewidths was unexpectedly introduced by a small amount of buffer gas in the cells. The buffer gas collisionally broadens the Lorentzian component of the Cs lineshape and shifts the resonant frequency. The amount of broadening is dependent on the gas species and its partial pressure inside the sample cell. An initial set of Cs cells, specified to be pure, was found to exhibit a Lorentzian linewidth to be substantially increased over the expected value. We note that similar observations have been recently reported [46].

In figure 19, the measured absorption coefficient (i.e. $-\ln(\mathcal{T}) = -\ln(\mathcal{T}_{meas}/\mathcal{B})$, where $\mathcal{T}_{meas}$ is the measured transmission data) is plotted on a logarithmic scale so that the Lorentzian pedestal of the Voigt function is visible (in a way similar to figure 1). Also shown for comparison is a Voigt profile of the expected Gaussian width that was consistent with the SPRT-measured temperature and Lorentzian width that was consistent with the instrumentally broadened Lorentzian linewidth. The fitted atomic Lorentzian linewidths were

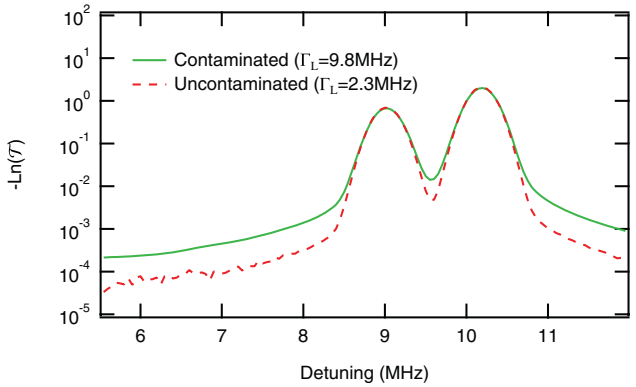

Figure 19. A plot of the measured absorption coefficient on a logarithmic scale, showing the Gaussian component at small detunings from the centres of resonance, and the Lorentzian pedestal. Plotted like this, it is easy to observe that the Lorentzian linewidth in the contaminated cell is the buffer-gas-broadened to 5 times that which is expected from the natural lifetime broadening alone.

$\Gamma_L = 9.73(1)$ MHz and $\Gamma_L = 9.88(1)$ MHz for Cells A and B at room temperature, respectively, which were ∼4 times greater than the expected value. The input power was also varied to ensure that this broadening was not an artefact of optical-pumping [47]. We note that we need a linearity and dynamic range of over 10^5:1 in absorption coefficient (i.e. 10^6:1 for a 10% deep absorption feature) to be able to see this effect shown on figure 19 with such clarity.

Table 7 shows the broadening and line centre shifts associated with some noble and nitrogen buffer gases. The parameters for each gas required to produce a total Lorentzian width of 9.5 MHz (composed of 2.3 MHz from the Cs natural linewidth, 0.05 MHz from the laser and 7.2 MHz from buffer gas broadening) is also shown and suggests that these cells were contaminated with a buffer gas pressure at approximately 1 torr.

Additionally, we used the spectrometer to compare the transition frequency of one of these contaminated cells with that of a replacement cell. The replacement cell showed no sign of gross contamination as evidenced by a measured Lorentzian linewidth that was consistent with the expected value. The line centres of both hyperfine transitions in the contaminated cell were shifted by 2.2(1) MHz relative to the non-broadened replacement cell. It is possible that a buffer gas mixture of 0.5 torr He and 1.9 torr of Ne was present.

In principle, these contaminated cells are completely valid experimental platforms for primary thermometry; the Gaussian width is unperturbed by the buffer gas. However, the unknown broadening means that the Lorentzian width in the function (6) must be left to be a free parameter in the fits. The retrieved Lorentzian value cannot be checked against the known values in the literature which which may reduce the confidence in the experiment. The contaminated cells were replaced with another two that were indeed uncontaminated and which generated a Lorentzian linewidth that was in agreement with the expected value. It is possible to make a very sensitive probe for possible cell contamination. The frequency shift of the absolute transition frequencies can be

Table 7. Buffer gas broadening and shift parameters at 300 K for some common gases [48].

Buffer Gas	HWHM Broadening (MHz/torr)	Shift (MHz/torr)	Broadening/ Shift	Pressure (torr)	Shift (MHz)
He	5	6.7	0.75	1.4	9.6
Ne	2.5	−2.9	−0.86	2.8	−8.3
Ar	5	−9	−0.56	1.4	−12
Kr	5	−2.7	−1.8	1.4	−3.9
Xe	5.5	−8	−0.69	1.3	−10
N_2	7.5	−7	−1.1	0.9	−6.7
'Cont.'	—	—	−3.4	—	−2.2
'Mix'	3.0	−0.9	−3.4	2.4	−2.2

Note: The pressure and associated centre frequency shift giving rise to a broadening of 7.2 MHz is also given for each gas. The measured parameters in the contaminated cell is indicated in the row labeled 'Cont.' Parameters for a buffer gas mixture ('Mix.') of 0.5 torr of He and 1.9 torr of Ne is also computed.

determined using an optical frequency comb stabilised to a primary frequency standard as a measure of gas purity. A measured shift of a Cs transition frequency shift of less than 1 kHz and a typical broadening-to-shift ratio of −0.5 (table 7), would produce a maximum perturbation to the Doppler width at the ∼3 ppm level if the Γ_L was fixed to the incorrect value.

7.2. Fundamental limits

Statistical: If the atomic lifetime is well known then we can fix its value in the atomic model—in this case, the only free atomic parameter in the model is the Doppler width. A simulations for conditions very similar to those used in this experiment (50 MHz frequency sampling over 6.5 GHz) shows that the statistical noise in $u(\Gamma_D^{fit}) \sim 0.33\sigma$ where σ is the standard deviation of the transmission noise divided by the amplitude of the absorption feature. If we leave both Γ_L and Γ_D free in the fitting algorithms then the statistical uncertainty contribution to $u(\Gamma_D^{fit})$ is around 5 times higher due to the strong correlation between Γ_L and Γ_D in a Voigt profile. Due to a lack of knowledge of the exact linewidth of the probing laser (see above), together with uncertainty in the literature value of Γ_L we did not fix Γ_L.

Experimentally, we estimate the uncertainty in Γ_D^{fit} at each power level by extracting Γ_D^{fit} from a large number of independent scans under the same conditions. We find a statistical error of 2.9 ppm in Γ_D^{fit}, (which represents 5.8 ppm in $u_r(k_B)$) averaged over all 1000 scans at a range of incident powers. We see that this is consistent with the simulation mentioned above. The uncertainty in the single-point transmission noise is ∼10 ppm (12.5 Hz bandwidth) at the highest incident powers, which is only a factor of 2 above the photon shot noise. After averaging over 200 scans at the same power we see $\sigma \sim 1$ ppm which predicts a $u(\Gamma_D^{fit}) \sim 2$ ppm. In order to lower this number (in a given integration time) we can use more power although this will necessarily increase the optical pumping effects (see below). Nonetheless, a 1 month measurement campaign at the current level of performance would yield a 0.1 ppm statistical noise. Alternatively, if we were able to fix Γ_L to a known value then a 0.1 ppm precision result would be obtainable after just 1 day of measurement.

Knowledge of Natural Lifetime: Given the result above it would be useful to fix Γ_L during the fitting procedure although this depends on its value being very well known. Under conditions in which $\Gamma_L < 0.5\Gamma_D$ we find that by fixing Γ_L to the wrong value (call this $\Delta\Gamma_L$) will translate into an error of Γ_D: $\Delta\Gamma_D \sim -1\Delta\Gamma_L$. The best known lifetime (to our knowledge) is that of the Cs 5P level [41] which gives a Lorentzian of width 2.287(6) MHz. This uncertainty in Γ_L delivers an uncertainty in Γ_D of 30 ppm. In our experiment an additional uncertainty on the linewidth of the probing laser (18 kHz) means that fixing the Γ_L is not sensible to obtain the best result. It is clear that were we to make use of a different atom and transition with a longer lived excited state then there is potential for the absolute error in the linewidth determination to be smaller and result in less uncertainty.

Instead, as mentioned above we have chosen to leave both parameters free in the fitting algorithm which allows us to obtain an independent measure for both Γ_L and Γ_D in the same experiment. In this case it is important to realise that our experiment is a means to making extremely high resolution measurements on atomic systems. Indeed, using our approach we are able to measure the f_{HFS} with an accuracy that exceeds that of the previous best determination in [49]. Saturation spectroscopy in vapour cells is appealing for its ease of implementation, but as Gerginov *et al* identified, such approaches often suffered from optical-pumping-related systematics that arise from attempting to form sub-Doppler absorption features. In this work we avoid this problem by using linear absorption, by greatly improving the detection sensitivity, and by extrapolating our results to the limit of zero probe intensity. The unknown linewidth of the probe laser prevents us from similarly generating an improved value for the lifetime of the Cs 5P level but were this to be improved then the experiment would already generate a better value than that known in the literature.

Were we to suppress etalons in the experiment to the 0.1 ppm level (as described above) and then make use of a longer measurement campaign (1 month) then we estimate an ability to measure the Γ_L to 300 Hz which would be the most precise measurement of an atomic lifetime. This in turn would allow a measurement of Γ_D to 1.5 ppm.

Vapour Self-Heating: One question that arises is whether the absorbed laser power could directly heat the thermal shield and/or the alkali gas itself. The beam size and intensity

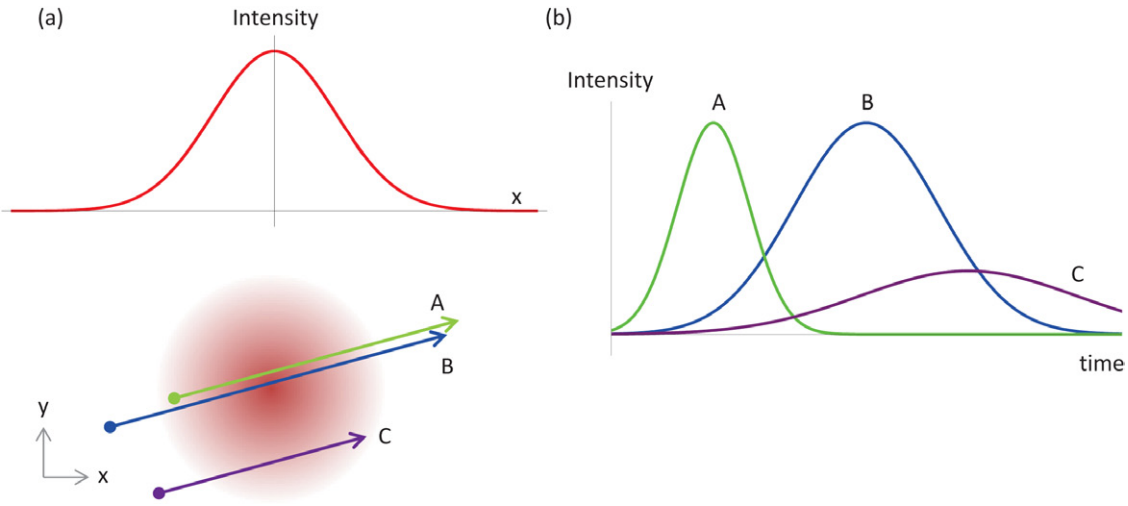

Figure 20. Schematic view of atoms flying through probe laser beam.

has been adjusted so that any particular atom is exceedingly unlikely to interact with more than one photon. Thus, the light doesn't cause any self-heating of the atomic vapour itself through random emission and absorption processes. This does leave room for some small net peturbation to the atomic velocities. We have directly considered this in [23] and find that the effective shift to Γ_D is below 1 ppb from this effect. On the other hand the scattered radiation may pass out of the glass cell and heat the thermal shield itself. The maximum scattered power is around 500 nW at the highest incident intensities which causes a heating of the copper heat shield (4.2 kg) over the scan time (30s) of around 80 nK or again less than 1 ppb fractional temperature change.

Optical Pumping: The image shown on figure 20 provides a more in-depth illustration of the interaction of a cw laser probe with a thermal vapour. Any particular atom will experience a variable duration interaction with the probe field with an additional strong variation in intensity depending on its transverse velocity and impact factor with the optical beam. When we combine this stochastic interaction with the potential of the atomic population to be optically pumped into the other ground state of the atom (which doesn't interact with the optical field) we find a modification to the atomic lineshape which means it is no longer a simple Voigt function. This has been modelled in depth in [25, 31].

We note that when we include this theory in our fit model then we can explain all of the features in the transmission measurement to the measurement noise floor (see top panel of figure 14) so there is no statistically significant signature of optical pumping. We find a maximum deviation of 15 ppm in $u_r(\Gamma_D^{\rm fit})$ at the *maximum* probe intensity associated with higher-order optical pumping corrections. Of course, these systematic shifts will be smaller at lower intensities. By forming the weighted average of all intensities we have used in the experiment contributes 7.5 ppm to $u_r(\Gamma_D^{\rm fit})$, (15 ppm in $u_r(k_B)$). This correction can be readily improved through three approaches: either a measurement at lower power than the current measurements (a reduction by a factor of 4 would reduce the correction by a factor of 16), a more careful measurement over a wider range of input powers would allow a better estimate of the sensitivity, or more measurements at any particular power would yield a more precise measurement of the shift. We estimate that a 1 month measurement campaign would yield a residual optical pumping correction of better than 1 ppm in $u_r(k_B)$.

8. Conclusion

We have provided an extensive description of the experimental details and limits associated with atomic Doppler Broadened Thermometry. We believe that many of these details are equally salient to other spectroscopic experiments and thus that they will be useful more broadly. We have obtained a value for Boltzmann's constant of $k_B = 1.380\,545(98) \times 10^{-23}$ J K^{-1}, with a precision of a 6 ppm in a few hour measurement and an accuracy limited by uncertainties in the Lorentzian component of the linewidth. We have outlined a path for improving these limits which would open the way to a ~2 ppm measurement using atomic spectroscopy.

Acknowledgments

The authors wish to thank the NIST Precision Measurement Grants Program, the Australian Research Council through the CE110001013, FT0991631 and DP1094500 grants for funding this work. The authors also wish to acknowledge the South Australian Government who have provided generous financial support through the Premier's Science and Research Fund. We thank M Moldover and G Strouse of NIST, and Mark Ballico from NMI Australia, for their contributions to the SPRT thermometry. We thank S Bize for his suggestion of the use of Hg as a potential atomic absorber in DBT.

References

[1] Fellmuth B, Gaiser C and Fischer J 2006 *Meas. Sci. Technol.* **17** R145
[2] Koch J A, Stewart R E, Beiersdorfer P, Shepherd R, Schneider M B, Miles A R, Scott H A, Smalyuk V A and Hsing W W 2012 *Rev. Sci. Instrum.* **83** 10E127

[3] Stewart G, Johnstone W, Bain J R P, Ruxton K and Duffin K 2011 *J. Lightwave Technol.* **29** 811
[4] Grefenstette B *et al* 2014 *Nature* **506** 339
[5] Hayato A *et al* 2010 *Astrophys. J.* **725** 894
[6] Wood C S, Bennett S C, Cho D, Masterson B P, Roberts J L, Tanner C E and Wieman C E 1997 *Science* **275** 1759
[7] Jönsson P, Ynnerman A, Froese Fischer C, Godefroid M R and Olsen J 1996 *Phys. Rev.* **53** 4021
[8] Brage T, Fischer C F and Jönsson P 1994 *Phys. Rev.* **49** 2181
[9] Bordé C 2005 *Phil. Trans. R. Soc.* **363** 2177
[10] Bordé C 2009 *C. R. Phys.* **10** 866
[11] Truong G W, Stace T M, May E F and Luiten A N 2015 To be submitted to Optica
[12] Moldover M, Trusler J, Edwards T, Mehl J and Davis R 1988 *Phys. Rev. Lett.* **60** 249
[13] Pitre L, Sparasci F, Truong D, Guillou A, Risegari L and Himbert M 2011 *Int. J. Thermophys.* **32** 1825
[14] de Podesta M, Underwood R, Sutton G, Morantz P, Harris P, Mark D F, Stuart F M, Vargha G and Machin G 2013 *Metrologia* **50** 354
[15] Cohen E R and DuMond J W M 1965 *Rev. Mod. Phys.* **37** 537
[16] Daussy C *et al* 2007 *Phys. Rev. Lett.* **98** 250801
[17] Casa G *et al* 2008 *Phys. Rev. Lett.* **100** 200801
[18] Triki M, Lemarchand C, Darquié B, Sow P L T, Roncin V, Chardonnet C and Daussy C 2012 *Phys. Rev.* **85** 062510
[19] Lemarchand C, Triki M, Darquié B, Sow P L T, Mejri S, Chardonnet C, Bordé C J and Daussy C 2012 *J. Phys.: Conf. Ser.* **397** 012028
[20] Cygan A, Lisak D, Trawiński R S and Ciurylo R 2010 *Phys. Rev.* **82** 032515
[21] Hald J, Nielsen L, Petersen J C, Varming P and Pedersen J E 2011 *Opt. Express* **19** 2052
[22] Gianfrani L 2012 *J. Phys.: Conf. Ser.* **397** 012029
[23] Truong G W, Anstie J D, May E F, Stace T M and Luiten A N 2012 *Phys. Rev.* **86** 030501
[24] Truong G W, May E F, Stace T M and Luiten A N 2011 *Phys. Rev.* **83** 033805
[25] Stace T M, Truong G W, Anstie J, May E F and Luiten A N 2012 *Phys. Rev.* **86** 012506
[26] Stace T M and Luiten A N 2010 *Phys. Rev.* **81** 033848
[27] Budker D http://budker.berkeley.edu/ADM/
[28] Anderson T, Magnuson J and Lucht R P 2007 *Appl. Phys.* **87** 341
[29] Alnis J, Gustafsson U, Somesfalean G and Svanberg S 2000 *Appl. Phys. Lett.* **76** 1234
[30] Magnuson J K, Anderson T N, Lucht R P, Vijayasarathy U A, Oh H, Annamalai K and Caton J A 2008 *Energy Fuels* **22** 3029
[31] Truong G W, Anstie J D, May E F, Stace T M and Luiten A N 2015 *Nat. Commun.* under review
[32] PDESolutions 2008 FlexPDE v.5.0.21
[33] Kittel C and Kroemer H 1980 *Thermal Physics* (San Francisco, CA: Freeman)
[34] Modest M 2003 *Radiative Heat Transfer* (Amsterdam: Elsevier)
[35] Truong G W 2014 Quantitative atomic gas spectroscopy for the determination of the Boltzmann constant and primary thermometry *PhD Thesis* The University of Western Australia
[36] Siegel R 2001 *Thermal Radiation Heat Transfer* 5th edn (London: Taylor and Francis)
[37] Preston-Thomas H 1990 *Metrologia* **27** 3
[38] Fischer J, Podesta M, Hill K, Moldover M, Pitre L, Rusby R, Steur P, Tamura O, White R and Wolber L 2011 *Int. J. Thermophys.* **32** 12
[39] ASL 2014 *Primary Standard Thermometry Bridges* (Surrey: ASL)
[40] Lemarchand C, Mejri S, Sow P L T, Triki M, Tokunaga S K, Briaudeau S, Chardonnet C, Darquié B and Daussy C 2013 *Metrologia* **50** 623
[41] Steck D (unpublished) (http://steck.us/alkalidata/)
[42] Stace T M 2010 *Phys. Rev.* **82** 011611
[43] Truong G W, Anstie J D, May E F, Stace T M and Luiten A N 2015 in preparation
[44] Domenico G D, Schilt S and Thomann P 2010 *Appl. Opt.* **49** 4801
[45] Jiang H, Kéfélian F, Lemonde P, Clairon A and Santarelli G 2010 *Opt. Express* **18** 16849
[46] Wu C M, Liu T W, Wu M H, Lee R K and Cheng W Y 2013 *Opt. Lett.* **38** 3186
[47] Truong G W, Anstie J D, May E F, Stace T M and Luiten A N 2012 *Phys. Rev.* **86** 030501
[48] Demtröder W 1981 *Laser Spectroscopy: Basic Concepts and Instrumentation* (Berlin: Springer)
[49] Gerginov V, Calkins K, Tanner C, McFerran J, Diddams S, Bartels A and Hollberg L 2006 *Phys. Rev.* **73** 032504