Learning from the time and length redefinitions, and the metre demotion

BY JOHN L. HALL*

JILA, University of Colorado and NIST, Boulder, CO 80309, USA

After discussing several issues in a future redefinition of the kilogram, this paper considers the lessons that one might have learned from the analogous redefinitions of the metre and the second. The progress of length metrology was slow and steady, from seven digits reproducibility with the 1889 X-shaped metre prototype, to nine digits with Kr lamps, to 11 digits with the 1983 redefinition of the metre using the speed of light. With laser cooling, the Cs clock improved to 15, now 16, digits (and so also astronomical distance measurements could improve). Laser-cooled ions, and now atoms captured and cooled in an optical lattice, enable accuracy capability of three different optical frequency references to exceed 17 digits, i.e. better than time itself. The optical comb and related techniques vastly simplify frequency comparisons. Such progress stimulates a new satellite experiment, the STAR Mission (Space–Time Asymmetry Research). The goal is to test at the 1E–18 level frequency shifts owing to spatial anisotropy, position, gravitational potential and boost. The onboard optical clock will use stabilization to a molecular transition in I2 or HCCH or CO2. The length etalons will be multiply redundant, with stability at the thermo-mechanical mirror motion limit. For a ULE glass etalon spacer (1987), I measure length creep approximately \(-1.5E–12/d\), i.e. below 1E–14 over the 500s satellite spin period.

Keywords: choosing base units; frequency stability; optical atomic clocks; defining the speed of light; redefinition of the kilogram

1. Background

With the great progress in the international ‘Avogadro project’, working to make the connection between atomic mass and the International System of Units (SI) mass scales, it will ultimately be appropriate to consider a redefinition of the SI kilogram unit [1]. The existing definition, in place for more than a century, speaks in terms of the mass of a particular metallic blob being kept in Sévres, France, by the BIPM (Bureau International des Poids et Mesures; http://www.bipm.org/utils/common/pdf/si_summary_en.pdf; http://www.bipm.org/utils/common/pdf/si_brochure_8_en.pdf; http://www.bipm.org/utils/common/pdf/24_CGPM_Convocation_Draft_ResoluUtion_A.pdf). The originators of this definition saw fit to have another approximately 40 such prototype reference masses produced and compared with the International Kilogram, item ‘K’ of the prototypes. Even if this example were to be suitable

*jhall@jila.colorado.edu

One contribution of 15 to a Discussion Meeting Issue ‘The new SI based on fundamental constants’.
Learning from earlier SI redefinitions

regarding its stability (which may be only marginally so), it does not match our objectives of having an international system of physical references which could be independently realized by any nation. Of course, it is natural to look to atomic physics, as that field has so amply supplied metrology specialists with atomic frequency, and two quantum relationships concerning electrical quantities [2]. Collectively, one can thus arrive at a reliable reference for time (for 40+ years the Cs hyperfine transition), and for electrical resistance (von Klitzing’s effect) and electrical potential (Josephson’s effect). These two latter quantum effects, in fact, are the basis for modern electrical standards, with reproducibility approximately two orders better than their connection with SI units, which is provided in the $3E^{-8}$ domain by the ‘watt balance’ experiments. So, in addition to the kg issue, the coming task of reworking the SI still also requires to fully rationalize the situation between the electrical units as maintained, and the existing SI. Also, as we will see, there are new inputs from time to time which may be directly of interest and utility with regard to reworking the SI. One example is the experimental measurement of the Avogadro number, which has recently been reported by an international collaboration. There are also auxiliary measurements such as the Rydberg and the fine structure constant, $\alpha$, which may provide constraints, as their theoretical interconnections are believed to be secure. Modern miniaturization of semiconductor components also provides us with a single electron turnstile transistor where the current is passed one charge at a time, yielding basically an ideal ability, in principle, to relate current to the number of charges and the time interval of their passage. However, it will still take some effort to develop the desired accuracy of mapping this nanoscopic current into the SI ampere. The three electrical quantum effects have been used in a triangle fashion to confirm the physics initially, and, with the enhanced confidence thus obtained, we have already pressed on to use these quantum phenomena to provide our electrical working standards, and through simple equations a route to the desired SI base standards.

But which exact scaling numbers are we to use? Is it acceptable for us to use ‘metrological constants’, combinations of physical constants such as $e$ and $h$ that are not precisely known in SI units? In fact, the developing consensus is to move to making the physics numbers be the adopted values, and calculate out what our mass, energy and temperature units will be. The thoughtful utilization of this large amount of data poses a very interesting task, deciding how best to provide low uncertainty in the base units. This work is somewhat similar to the task involved in the periodic analyses undertaken by CODATA (for the benefit of readers from other fields, a list of abbreviations used in this article is given in table 1), where the objective is to provide the best values of the physical constants, especially including the effort needed to make the results consistent—to within the assigned errors—with the known physics relationships. However, the current task of reworking the SI is substantially larger, because high precision for some base quantities is of higher significance than for others. For example, I would not object if the uncertainty in the volt could be reduced somewhat by accepting a small corresponding degradation in the definition of the kg. In the user community, there will be many opinions, and probably some clear concerns will emerge. I urge us to proceed deliberately and discuss our proposed changes widely before the changes are made.

Phil. Trans. R. Soc. A (2011)
Table 1. Abbreviations used in this paper.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMU</td>
<td>atomic mass unit (symbol, u) or dalton (symbol, Da) is the unit of mass on an atomic or molecular scale. It is defined as one-twelfth of the rest mass of an unbound atom of carbon-12</td>
</tr>
<tr>
<td>CGPM</td>
<td>the General Conference on Weights and Measures is the English name of the Conférence générale des poids et mesures (CGPM, never GCWM). It is one of the three organizations established to maintain the International System of Units (SI) under the terms of the Convention du Mètre (Metre Convention) of 1875</td>
</tr>
<tr>
<td>CODATA</td>
<td>the Committee on Data for Science and Technology; an interdisciplinary Scientific Committee of the International Council for Science (ICSU) established 40 years ago</td>
</tr>
<tr>
<td>CW</td>
<td>this refers to operation of an electronic/optical source on a continuous wave basis, rather than a pulsed basis</td>
</tr>
<tr>
<td>EMF</td>
<td>electromotive force, measured in volts of electrical potential difference</td>
</tr>
<tr>
<td>IR</td>
<td>infrared, electromagnetic radiation with wavelengths longer than those of visible light</td>
</tr>
<tr>
<td>rf</td>
<td>radio frequency</td>
</tr>
<tr>
<td>ULE</td>
<td>Corning Code 7972 Ultra Low Expansion Glass is a titania silicate glass with unique low thermal expansion characteristics. Mention is for scientific communication purposes alone</td>
</tr>
</tbody>
</table>

But do let us enter a bit into this discussion. On the atomic scale, to our knowledge, atomic masses also have the delicious character of universality. Very high precision is attainable with ion traps in comparing atomic, more properly ionic, masses, and 10 and 11 digits are achievable for their ratios [3,4]. However, there is still the problem of knowing the magnetic field if we wish to know the moments and masses in SI units. We can know the current with reference to quantum resistance and quantum voltage standards, but, to define magnetic field, we need to know some geometry about the coil and the current distributions. Even though the metre is strictly defined from an optical frequency, and some heroic interferometric experiments [5] have produced 10 digits (or better) knowledge of the effective separation of two suitably curved spherical mirror surfaces, it seems daunting to adequately define the location of the current-carrying wires or even to define the outer surface of a cylindrical coil form. The concept of working out the positions of the multi-layers of fine current-carrying wires is best ‘left as an exercise for the student’. So we can expect only high precision ratios of atomic masses.

If we wish to actually try to compare the gravitational force on a mass to an electrically generated force, it is attractive to use a macroscopic mass. Yes, it can be supported by electromagnets and the currents can be measured accurately, but the coil geometry problem seemed permanent. The ingenious idea of Brian Kibble was to deliberately move the two coils relative to each other, and measure the generated EMF to calibrate the coil geometry. This is the core of the watt balance approach to mass measurement [6], using the coils with current to make a force and then calibrating the geometry of the same coils. But, despite extended efforts, this approach has not yet surpassed the eight-digit level that can be approximated by precision weighings. Still, the deduction of Planck’s constant to $5E^{-8}$ has been possible in this manner [7]. So we seem to be getting closer to the objective—the possibility of a reworked ‘SI’ based on atomic physics, without any artefact standards.

*Phil. Trans. R. Soc. A* (2011)
We all can—and should—enthusiastically thank the colleagues of the ‘Avogadro project’, who are diligently working to fabricate a 1kg mass of technically adequate single-crystal silicon [1]. Earlier experiments along this highway made clear the need for mono-isotope material, while the electronics community has already learned how to approximate the dislocation-free, high-purity, controlled vacancy, single-crystal material that would be the natural conceptual requirement. It seems likely that ways can be found to tolerate small deviations from these ideality conditions. Indeed, the present limitations seem to come from surface definition issues, with perhaps some of the optical polishing compounds becoming embedded in the surface layers. Of course, there will be an oxide coating on the silicon sphere, so it will probably be a good plan to strip the porous natural oxide layer and replace it with a deliberate high-density oxide layer. We then need to measure the added layer’s thickness through ellipsometry or other methods to be able to guess the added weight. And finally, probably a half-decade hence, we can hope that the present 3E–8 uncertainty of the Avogadro constant can very likely be reduced threefold more, offering us an eight-digit knowledge of the number of silicon atoms contained in the 1kg sphere, which perhaps would represent an adequate connection between the atomic mass of $^{28}\text{Si}$ and the existing kg SI unit.

We would add that the $^{28}\text{Si}$ mass in AMU was recently measured to 2E–11 [4]. At this point, we have wonderful precision within the atomic-scaled units, but expressing the mass of a single $^{28}\text{Si}$ atom—or the AMU itself—by means of even an ideal Avogadro experiment would carry an undesirable uncertainty from the present kg realization. Of course, in whatever way we will choose to go forward with the new unit system, we wish to have the new quantities realizable with improved accuracy, while retaining the ability to know the macroscopic laboratory masses with an uncertainty not much more than the present kg uncertainty. Suppose that this vision leads one to consider adopting a fixed value for Avogadro’s number also. Then perhaps we have settled the mass question? The answer is NO! Because the many links between atomic quantities are known to much higher accuracy than is achieved in weighing the macroscopic masses. For example, immediately one would remember that the dimensions of Planck’s constant are joules seconds, and joules are kg m$^2$ s$^{-2}$. Of course $h$ or $\hbar$ appears in the basic expression of the quantum electrical quantities, and so the uncertainty which began with the kg could easily appear in the units of these remarkable electrical relationships, which may be exact in theory. This feels like a seriously poor highway to follow, as the electrical quantities are already measurable more precisely than the laboratory masses. Furthermore, I suggest that we are more likely to find further delicious surprises from nature among the atomic and quantum worlds, rather than the macro-world of kilograms or tonnes.

So what should we choose to define [2]? Certainly, it makes sense to make the new SI units rely on the physics of quantum mechanics: we have already chosen this road for the second and its secondary quantity, the metre. On the electrical side, the single electron transistor clearly serves to produce a digital relationship between current and charge, so between the ampere, the coulomb and the frequency. The Planck constant appears in both of the remaining electrical relationships, which I remind are the most precise relationships we have ever discovered among variables we had supposed to be continuous. The Josephson effect provides an essentially ideal frequency–voltage mapping, with
$eV = n\hbar\omega/2$, written as $(\omega/2\pi) = K_J \cdot V/n$, using the Josephson constant $K_J = 2e/\hbar$ approximately $483.6$ THz $V^{-1}$. Von Klitzing discovered the integer quantum Hall effect, which provides plateaus of quantum physics-defined electrical resistance, according to $R_k = \hbar/e^2$, about $25.8$ k$\Omega$. So, it seems a rather unique path that leads to the proposal that it is the Planck constant which is to be redefined, which will bring the quantum relationships involving resistance and voltage into the SI [8]. Mass will then be a measured quantity, in terms of atomic physics parameters. I expect little damage will be done thereby, as the measurements that lead to a mass determination of laboratory scale are challenged by a plethora of details that lead to imprecision and perhaps errors. Probably, realizing the derived unit kg at the 100 ppb level, or ideally 10-fold below, would be sufficient.

But choosing the ‘best’ value for $\hbar$ is challenging since, as noted earlier, interesting constraints arise from other quantum physics relationships. For example, the (inverse) fine structure constant $\alpha^{-1}$ is about $137.036$, and is known theoretically to be $4\pi \varepsilon_0\hbar c/e^2$. With laser methods, it has been measured in hydrogen spectroscopy to some seven parts in $10^{10}$. Similarly, experiments have yielded the Rydberg spectroscopic constant $R_\infty = \alpha^2 m_e c^2/2\hbar$ (approx. $10.973 \times 10^6$ m$^{-1}$) to within 7 parts in $10^{12}$! These and many more considerations regarding the merger of slightly divergent data, with differing uncertainties, into a coherent set of ‘best values’ form a fascinating art, and the interested reader is referred to the recent CODATA reports [9].

I now turn to the fact that the BIPM and metrology community has experience that is relevant to the question of ‘when should we undertake the redefinition of the kg by means of the fixing of Planck’s constant?’ (This change would almost surely also include rationalization of the electrical units.) Let me state my conclusion early so that the reader does not spend energy in checking this author for possible bias. This is the right path, but we must wait a bit until the time to change is right, which I suppose may not be until a half-dozen years from now. Or perhaps, at the 2015 CGPM meeting? So let us turn to the reasons behind this opinion.

First of all, we need a convenient way to have a secondary representation of the kg in any laboratory. The basic requirement before a redefinition occurs is that no stakeholders are ‘technically frozen-out’ or ‘priced-out’ of their access to the realization of the base SI unit for their country. An equivalent statement is that the apparatus for such realizations be well developed, mature and even available commercially if possible at the time of the change.

What is still needed? We really ought to know what happens in the next iteration of the silicon Avogadro measurement—that would be a minimum expression of our respect for the tremendous efforts already invested by this international collaboration. Furthermore, the NPL watt balance apparatus has been transferred to the NRC, and one can expect those colleagues will jump on the opportunity with enthusiasm. One really should expect and must request further progress on this highway, as well as on the Avogadro road. If the community of metrology executives of the CGPM were to act officially now, and probably be forced to issue a correction 5 years later, this would be absurd: it would be profoundly insulting to the large number of working-level colleagues who have invested a large part, perhaps a major part, of their working careers into these measurements and comparisons. This error would be crushing to the spirits of the
participating national laboratories, and would hardly help them in their annual challenge to obtain adequate domestic support for infrastructure-type research. In almost all countries, it is just not a popularly recognized need to have a new way to define the kg. This is in contrast with the Josephson and the von Klitzing electrical conversion constants that have been widely tested and confirmed, and no users are prevented from doing their science because these conversion units have not yet been formalized into the SI. When the redefinition does occur, it seems basically impossible that these conversion values would be changed slightly to better conserve the kg unit. So, in my view, it really does not matter if we wait until we are actually ready to take this ‘new SI’ step.

Secondly, a prompt redefinition invites full public disaster in connection with the traceability of the mass unit. Now each nation is responsible for checking the submitted secondary reference kilogram standards, which it can do as most of the strongly industrialized countries still have their International Prototype kilogram mass standard from 1889. They have been checked several times at the BIPM and the deviations noted. However, today the circumstances are that even a consortium of national laboratories has not yet completely worked out all the problems in connecting the atomic and macroscopic worlds via silicon at the desired 10 ppb level [1]. Several major efforts via the watt balance highway also have yet to reach the desired accuracy [6,7]. So, with an atomic physics definition as the only route to the kg, it surely would be too much to expect that there is now an accessible, reliable and solid path for even the less wealthy countries. In my opinion, taking such a redefinition prematurely would simply be arrogant and disrespectful to all the practitioners in the field. I fail to see the benefit.

So I believe some thought is needed, additional patience is needed, new experiments should be invented and new infrastructure capability needs to be developed, so that any nation can have a cost-effective route to apply the new mass definition in their commerce and technology.

In the interim, one wonders if this nirvana could be effectively approached by making isotope-enriched Si spheres in a much larger quantity, and in a smaller mass, perhaps 0.3 or 0.1 kg. Then a little precision will be lost in the actual weighings, but the problem is not in the measuring process, but rather in the quality of the reference masses. Mono-isotope dislocation-free single-crystal silicon ping-pong balls might be about right. Another topic connected to a slightly smaller size scale is that the polishing and length metrology begin to run out when one is discussing the last few atomic planes. So it is slightly worse, but only as $(M)^{-1/3}$, if the mass is decreased. Under this scenario, any nation could choose to accept the Mass Calibration Certificate for their specimens, or could invest in determining the masses by an Avogadro experiment. But, without a new scheme, the technical difficulty of a single nation contributing effectively to the Avogadro seems daunting. While awaiting a new physics inspiration or further advances by this collaboration team towards the silicon kg, one practical step would be in exploring the limitations around using certified kilogram secondary standards, such as those available in cost-effective stainless steel.

An important argument was presented by Dr Terry Quinn, which I reproduce here with my own biases probably showing. The mass artefact has the property that it can be compared with another mass specimen with a precision which is substantially better than the estimated long-term stability of the kilogram prototypes, say over 10 years. It appears at least plausible that the kg has become
J. L. Hall

perhaps is always becoming?) lighter over time, as the weighing results for the several national prototypes seem on average to be trending towards higher apparent mass. Perhaps a little material is removed from the kg each time by the cleaning? So even though the watt balance approach has not yet reached the accuracy level that we would wish for, it can perform an extremely important function by bringing clearly to our attention any long-term drift of the employed prototypes. I do know that a large amount of money changes hands on the basis of weight of the purchased materials, and perhaps grains such as wheat, corn and rice may have the largest dollar value. But the moisture content of these natural materials is by far not standardized, and I expect that any problems with the base unit are of a vastly less significant scale.

2. Can we learn from the last redefinitions within the International System of Units?

As a potential aid to experientially based decisions regarding when for the mass decision, I invite the reader to reflect on the experiences that our community has had when the time and length redefinitions occurred. As for the time redefinition in 1960, just please remember that this was when a base frequency of one cycle per year was advanced by experts from one particular branch of science. This standard (of inverse frequency) was then chosen by the CGPM, and, while it was perhaps philosophically valuable in view of its expected stability, the redefinition took no account of the needs of the huge industrial base of people working in the tens of gigahertz with advanced radar and communication systems. I expect the most charitable statement is simply that a mistake was made. As documentation of this outlook, it was only 7 years later that their next choice was the caesium atomic clock, which had become adequately mature during that interval. This Cs atom choice has been a true winner, with continuing improvements aggregating to about 1 million-fold over the 60 years of experience since that Cs definition of time. In part, these huge improvements were possible because the technology was new, so a first realization was not likely to be near the optimum. Additionally, the statement of the definition was highly conceptualized and free of detail, leaving room for the great innovations that have occurred, especially using laser-cooled atoms in Cs atomic fountain clocks.

On the length metrology side, progress was slow and steady, beginning from the approximately seven digits reproducibility with the 1889 X-shaped metre prototype, to nine digits with Kr lamps, following the 1960 redefinition. When laser interferometry became possible and then progressed to 10 digits, the asymmetry of the krypton line was transparently observable. With the 1983 redefinition of the metre using the speed of light, we locked in the existing metre uncertainty of \(4 \times 10^{-9}\) into all previous length measurements. With laser cooling, the Cs clock improved to 15, and now 16, digits. So also astronomical distance measurements could improve. But with laboratory length measurements, even with lasers, there seems to be an asymptote near 11 digits, basically owing to the hard-to-characterize phase shifts associated with imperfect surface properties [5,10].

More recently, cold atoms can be captured and further cooled in an optical lattice, augmenting the use of laser-cooled trapped ions, with both technologies enabling accuracy capability of three different optical frequency references to be
Learning from earlier SI redefinitions

now entering the 17th digit ($\text{Hg}^+ [11]; \text{Al}^+ [12]; \text{Sr} [13]$). Four more systems ($\text{Hg} 1.13 \text{PHz}; \text{Yb} 518 \text{THz}; \text{Yb}^+ 688 \text{THz}; \text{Yb}^{2+} 642 \text{THz}$) are rapidly approaching this level. This is the first occasion when any measurement capability exceeded that of time, which itself has always been our most precise reference [14]. The Laser Revolution, even after its 50th year, continues to provide powerful new tools for metrology. The optical comb and related techniques have vastly simplified frequency comparisons, with demonstrated accuracy capability to 19 digits [15]. The impact of the comb was something analogous to a phase transition of melting problems away from their intrinsic frozen state: it is obviously nearly impossible to measure across a million-fold frequency gap in a convenient and phase-coherent way, unless of course you happen to have a source with such a huge nonlinearity that you can jump the frequency gap in a single step!

3. Observations about the length redefinition (by a participant)

Let me just add a few words about the remembered experiences of a young enthusiastic booster of the significance of the new-fangled stable laser, this first one based on intra-cavity saturated absorption of free-flying methane molecules by the HeNe laser operating at 3392 nm (88.376 THz). By the end of 1971, my colleague Dick Barger and I had built several of these systems and all our laser colleagues were impressed to find how well they worked (figure 1).

The stability was approximately $10^{-13}$ (or below) by 100s and the reproducibility was about $10^{-11}$ when one changed the mirror alignment or power or whatever. By the end of 1972, we had a careful wavelength measurement against the krypton lamp and were delighted to contemplate what a boon this was going to bring to metrology people.

Well, as it works out, the world’s main length calibration activity was trying to calibrate gauge blocks or other etalons using the krypton lamp, introduced in 1960 to replace the metre bar as the legal definition, with the cadmium lamp as the actual working secondary standard. The 1 ppm gauge-block calibration

**Figure 1.** Measuring laser stability. Richard L. Barger (right) and J.L.H. (left) measuring relative frequency stability and reproducibility of an early generation of HeNe/methane-stabilized laser at 3392 nm (approx. 1971). Use of an offset-locked auxiliary laser (middle) allows continuous servo-locked tuning of a test laser (back right) through the resonance centre frequency. (Online version in colour.)
Figure 2. Getting length accuracy to where it is needed. The (rectangular) gauge block can transfer laboratory calibration standards to the workshop bench, at approximately 1 ppm. But the micrometre, even in skilled hands, can compare two objects with an uncertainty of just some few micrometres, which is about 10-fold worse. This is just sufficient for maintaining modern diesel and petrol engines, while the length definition is 10 orders more precise. (Online version in colour.)

The goal was equivalent to approximately 0.1 visible fringe, which could be obtained using the several different radiations from the krypton discharge. The gauge-block calibrations had customers who were calibrating precision micrometres for precision machine shops, where a few micrometres accuracy on a few centimetres object was their objective. Perhaps 10 ppm, only! By now, valve lifters and piston pins of modern turbocharged car engines actually can profit from tighter tolerances (figure 2).

These length colleagues were trying to measure optical fringe shifts, and had to deal with the generous thermal expansion of the gauge-block material. In 10 cm, one had more than a hundred thousand fringes, and needed only a reliable fringe division of 1/10 to support the desired 1 μm precision. It was not too troublesome for them even if there was too much N₂ slush-ice floating through the optical path in the krypton lamp’s cryostat. And there I was in 1971, offering them a fantastically stable—but invisible—laser beam for their task. At the time, there were probably not any real IR cameras or IR viewers in existence, and most certainly none were available outside some military night-vision laboratory anyway. Even now such IR viewers for 3400 nm are not so widespread (while for some reason there are huge amounts of tools available at 1500 nm!). So the 1971 CCDM meeting accepted my data, but not my enthusiasm for a quick change from the easily visible lines of Kr discharge. (Other lamps, such as Cd, were also in wide use.)

By contrast, my colleagues and I were measuring (small) frequency differences and it was absolutely easy. Dick Barger and I worked hard to refine our wavelength measurement [16] and so came to have to deal with the invisibility of this 3392 nm
Learning from earlier SI redefinitions

Figure 3. Comparing krypton and laser wavelengths. Fringes of 3392 nm laser (top) and 605.7 nm Kr standard lamp (middle trace). Bottom trace shows linear length sweep under frequency servo control. Interferometer is optimized for the Kr radiation. Spacing is 18 cm, so nominal Kr precision of $4 \times 10^{-9}$ can be reached in about 300 s averaging. But line asymmetry requires many such measurements versus aperture and focusing to obtain corrections to the ‘measured’ wavelength.

reference. Of course, it is straightforward to co-align a visible HeNe laser with the same confocal mode parameters, so as to facilitate the easy alignment of the IR optics. We soon discovered the reason that people such as Dick Rowley (NPL) and Pierre Giacomo (BIPM) were concerned with the symmetry of the 605.7 nm Kr line. It simply is that the wavelength of the Kr line was variable across the capillary aperture, which retrospectively is not surprising, considering that atoms near the capillary tube will have a lower net axial velocity under the influence of the DC axial electric field of the discharge. The light emission is from a high-lying Kr energy level, which is typically populated by recombination of an electron with the discharge-accelerated ion. Taking into account that each National Metrology Institute had a differing optical arrangement, the weighting of the light emission across the aperture was different in each set-up. Eventually, it was shown that the krypton metre was realizable at about $4 \times 10^{-9}$, after the several-fold larger systematic shift was suppressed by specification of the optical configuration and the line-splitting algorithm [17] (figure 3).

Let me make explicit the parallel with the present SI redefinition issue: what we offered then was the higher precision (in the IR) which was already available to all high-skill workers in the field, who would naturally share our conviction that the frequency of the HeNe/CH$_4$ standard was not going to change. It was known then in SI units with an uncertainty of ‘only’ $5 \times 10^{-10}$, but could be relied upon securely to ‘keep’ any measurements almost 100-fold better, even though they could not be expressed in SI Hz to the full precision limit. This feels to me like a close preview of the present situation where the quantum electrical units can be reproduced and disseminated to many digits beyond which they are known in SI units. But it really does not matter, as the conversion constants will surely not be changed numerically a few years hence when the redefinitions have become ripe for action.
In the laser case, it was not until some years later that it became useful to try to measure visible frequencies. For one thing, by then many laboratories had made combined wavelength and frequency measurements to confirm the basic soundness of the adopted speed of light number [18]. Secondly, we were all anxious for the metre redefinition to occur, but the colleagues who actually were performing metre-traceable measurements understandably felt that the frequency of a convenient visible reference laser (such as the 633 nm iodine-stabilized one) must be known first. A third factor was the wish for an advanced technology, better suited for connecting into the visible domain than were the point-contact diodes. This advance came in 1976 with the invention of a frequency-adding scheme by Chebotayev et al. [19] in a neon discharge—an elegant realization of phase-matched four-wave mixing. Accepting this idea, we were tasking ourselves to input the three different (but measurable!) laser colours that step-wise matched the neon transitions, which would lead to coherent population of the 3s2 state. This state does not even have to be inverted relative to the initial 2p4 level, as the emitted coherent red beam at 633 nm is radiated by the coherently emitting dipoles. This began the second long collaboration in Boulder, CO, with Ken Evenson’s frequency-measuring team, and led eventually in 1983 to a measurement of the red 633 nm iodine-stabilized HeNe laser in terms of the 3392 nm CH4 reference (and thereby gave the red laser’s emission in hertz) [20]. Similar kinds of frequency chains were built by T. Blaney, Gordon Edwards, David Knight and colleagues at the NPL. It was a huge pleasure to become acquainted with these NPL team members, who also furnished wavelength, frequency and speed of light numbers to the CCDM [21,22]. The NIST-Gaithersburg team had an interferometric comparison between red and IR [10], as well as nice stabilization results using isotope 129 molecular iodine. Bay and Luther at NIST made the first all-visible interferometric speed of light measurement [23]. The PTB offered precision measurements on \(^{129}\text{I}_2\) at 633 nm, as well as \(^{39}\text{Ar}^+\) laser stabilization. Our Japanese NRLM colleagues also contributed 633 nm wavelength measurements. The NRC colleagues provided a significant confirmation of the adopted \(c = 299792458\) number also [18], using an innovative chain based on several different CO2 lasers and their various harmonics and differences to multiply up coherently from microwaves, ultimately up into the visible [24]. Their CO2 wavelength measurement, along with the one at the MIT by Javan’s group [25], gave two more confirming values for \(c\). As was often the case, Chebotayev was the first to discuss a far-future possibility, an optical time scale [26].

So, by 1983, there were multiple and consistent frequency and wavelength measurements both for the 3392 nm methane-stabilized laser and for the 633 nm I2-stabilized laser. We were well positioned to recommend to the CGPM that the time was ready for the metre to be redefined. They even followed our CCDM’s recommendation, using the speed of light prescription. The outcome of such worldwide collaboration and long successful preparations by many laboratories led to the easy adoption of the speed of light, and even the effective abolition of the former namesake for the metric SI, its focal basis element of 100 years’ cooperation between almost 50 nations, the metre. Big changes can come with good preparations.

To summarize, this red laser measurement programme did not really begin immediately after the 1971 CCDM meeting, where the speed of light measurement was presented because the technology was not ready. I find an awesome
correspondence to the present conditions relating to defining $e$, $h$ and $N_A$ if a redefinition decision were made now, forcing the world to scramble to then operationally replace the $kg$. In my view, it is better if we pursue this change to the SI when we are actually ready. I look forward to joining the celebration of a renewed SI, when that time also is ripe!

One thing which has propelled this progress is the wide interest in frequency measurement as being the good highway towards precision and accuracy in measurement. This capability is leading to ever better optical frequency references and, as noted earlier, now at least three optical atomic systems have published results that are superior to the best results obtained with Cs atoms, that is to say the optical frequency results can now be more reproducible than time itself. What interesting things will we find [11]?

It was my good luck to be on duty when a number of concepts and technologies were simultaneously ready in 1999, allowing one to try out the frequency comb idea with light. The frequency comb concept was in use in the radio frequency (rf) days, and an effective rf comb-based frequency calibration system was even available in the 1940s, carrying the US Army’s name of BC-221. Baklanov & Chebotayev [27] and Eckstein et al. [28] saw the possibilities for the optical domain, even in the 1978 timeframe. A train of equally spaced repetitive pulses would indeed have interesting spectral properties, but to me it seemed to be like a discussion about whether angels would pack better on a round or hexagonal-shaped head of a pin. We just did not have any of those necessary optical things at that time. Then, in 1993, femtosecond-scale pulses became available with Ti:sapphire lasers [29]. I was lucky 6 years later to see a post-deadline paper at the 1999 CLEO meeting where a rainbow of colours was reported, generated at a 100 MHz rate using a mode-locked laser blasting its pulses through a specially designed multi-structured optical fibre [30]. The phase-locking that is forced by such remarkable nonlinearity of response ensures that the comb is strictly phase-locked internally, thus forming a basically ideal tool for the metrologist (figure 4).

Careful comparisons have been made between combs based on different lasers, using different fibre materials—changing everything that can be changed. None have ever revealed any deviation from the simple model. These important tests began in Ted Haensch’s laboratories. One such circle of measurements set a limit of $10^{-19}$ for the possible inaccuracy of the internal relationships within the comb [15]. By now, good comb systems are commercially available and make a perfect response to the question of how we will be able to make precision frequency measurements at this or that unconventional wavelength? We will be able. Trust me on this.

Our rf–optical frequency ‘multiplication’ factor can be some millions. But the factor between the atomic mass and the needed kilogram scale is more like $3 \times 10^{25}$-fold. How will we ever do this? Impossible? But just consider this: I was not sure the harmonic relationship could ever be found, when we had only 9.2 GHz and 88 THz to compare. The next issue was that the multiplied phase-noise would probably consume all the carrier frequency power into noise sidebands. Still, when the optical frequency comb arrived, it became easy to go 10 times and more beyond the objective of our previous impossible dream. New ideas, new dreams and good meetings stimulate people and open the new highways.
Figure 4. The optical frequency comb. Illustration shows optical frequency reference laser field (top wave) and the 1 GHz-rate short-pulse ‘stroboscope’ laser. These pulses are so short that they contain only a few optical cycles, so the superposition with the CW reference laser beam gives rather strong heterodyne beats, which is helpful for the rf beat measurement. In this illustration, the two pulses have the same shape, so carrier-envelope phase-shift is the same for the two successive pulses. Consequently, the optical frequency, seen to be stable relative to the pulses, is actually known to be an exact harmonic of the pulse rf. In practice, one often chooses to operate with stable rf offsets (rather than zero frequency) in the interest of better S/N. Illustration courtesy of Scott Diddams. (Online version in colour.)

4. Stable laser limitations

Of course, it is useful to have a laser frequency source that has an ultimately narrow linewidth. The rf FM methods allow such high S/N performance that the laser’s frequency noise is essentially replaced by the length noise of a high-finesse reference interferometer. The cavity drift initially is the main problem, so it has become standard to make this interferometer from ultra-low expansion glass. The Corning ULE material has a temperature zone somewhat below room temperature where the length variation is only quadratic with temperature variations, with a fractional rate typically $dL/L$ approximately $2 \times 10^{-9}$ ($dT/T)^2$. We can readily stabilize to 1 mK, and set to near the shortest length’s critical temperature to say 30 mK. This gives a length stability approximately $6 \times 10^{-9}$, or about 3 Hz. There is also a long-term creep of the cavity length, getting always shorter, which may be a dozen hertz per second of the optical frequency drift soon after the mirrors have been optically contacted to the spacer and the system mounted in vacuum. Within a year or so 1 Hz s$^{-1}$ should be reached. My ULE interferometer that was set up in 1987 has a drift rate of approximately 6 mHz s$^{-1}$ at present.

However, laboratory vibration can play bad tricks against our stability dreams. The acceleration forces do not bend things very much, say $dL/L$ approximately $2 \times 10^{-9}/g$. But laboratory floor vibrations can approach 1 milli-g, so we can expect frequency noise approximately $2 \times 10^{-12}$, or about 1 kHz r.m.s. at the vibration peak. Thus, vibration isolation is interesting, and factors of 10 or 30 are realistic.
Learning from earlier SI redefinitions

Figure 5. Vertical cavity design for reduced acceleration sensitivity of its length. The vertical mounting from below is via ‘wobbly’ Teflon posts that fit into the pockets shown at the bottom of the central flange. A small asymmetry of the mounting height in the central flange is calculated to minimize the cavity length change under vertical acceleration. Just by the fabrication precision, the sensitivity of this 70 mm long cavity is about 30 kHz ms$^{-2}$. This is low enough to give good results as shown in figure 6, but is only 12-fold below the sensitivity that would come from mounting at either end. Better vibration suppression factors were readily obtained with cylindrical spacers, so we speculate that the tapering towards the ends may not be adequately symmetrical. This cavity could be trimmed by adding small bits of In wire, but the operative limitation is thermal vibrations; see text and figure 6. (Online version in colour.)

Then it is time to redesign the cavity structure itself so as to use a natural symmetry to be more vibration-insensitive. One simple way is just to mount the cavity symmetrically, applying the support forces just at the cavity midplane. See figure 5 for a representative design.

Figure 6 shows the beat between two diode laser systems locked to two such independent cavities, located on separate tables and connected by a noise-cancelled optical fibre [31]. Apart from some problem with feedthrough of the air temperature cycling, this 1E–15 performance is rather satisfying. The beat can be recorded as approximately 300 mHz wide, but broadens to 2 Hz for longer averaging. The reference cavities are on small vibration isolation platforms on ordinary optical tables in the laboratory. The frequency noise rises for shorter...
times, but still amounts to much less than a radian phase excursion. However, the $1/f$ decrease rate could be associated with direct averaging over some small periodic perturbation. Perhaps it would be interesting to try a different optical power level to check where the shot-noise level is. However, see below.

One well-known effect has become visible and troubling with the current-generation ultrastable lasers, which transfer the passive stability of a high-finesse reference interferometer to the probe laser’s frequency. It is a little funny that the limitation on this Einstein’s Light Clock idea actually comes from something like Brownian motion, which he also studied in 1905. Numata et al. [32] have shown that an interesting level of r.m.s. position noise of the mirror surface is driven by thermal agitation. From the mirror substrate alone, the scale is approximately $4 \times 10^{-17}$ m/√Hz at 1 Hz Fourier frequency [32]. The laboratory vibration-resistant vertical cavities may be 70 mm long, and so we can essentially predict a spectral frequency density of 0.25 Hz/√Hz, with $1/\sqrt{f}$ frequency scaling. We show these dependences graphically in figure 7, along with data for the two best reference cavity systems studied up to now. Expressing the results in an Allen Deviation graph, such as figure 6, the thermal noise turns out to be flat at longer times. In the figure, the solid horizontal line is the thermal noise calculated for this cavity. So even coherent lasers are limited by thermal noise!

5. Other applications for standards’ technologies

There is a potential weakness of incorporating experimental facts, whose validity we can know only approximately, into statements that we come to rely on. Take for instance: ‘the Speed of Light is a Constant of Nature’. Well yes, let us do start with that and then see in which decimal the deviation will be found. We hardly ever find out something which is exact. Our statements usually have to say ‘in the
Learning from earlier SI redefinitions

Figure 7. Calculated thermal noise-generated frequency deviation spectral density, compared with the world’s best experimental results for laser stabilization. A longer cavity would reduce the frequency variations resulting from the mechanical motion, but would be more sensitive to laboratory vibrations. Integrating from a short time (high Fourier frequency) to longer times, the thermally generated FM noise sidebands have used up 50% of the carrier power by \( f \sim 0.3 \text{Hz} \). See Numata et al. [32] for details. Figure courtesy of Kenji Numata. (Online version in colour.)

non-relativistic approximation’, or ‘at low field strengths’, or invoke some other constraint. But a strictly constant speed of light, independent of gravitational field, direction, epoch, wavelength? It is a testable dream.

I sincerely hope you also are anxious when we so quickly abbreviate statements like: ‘We have tested the symmetry postulate with oxygen molecules and to the xxx sensitivity of our apparatus we find no deviation’. In the (physics) newspaper it says ‘scientists prove the validity of the symmetry postulate’. Feel free to substitute your own idealizations and examples. At some level you might think all these ‘exact’ symmetry things would break down. At least, I do. Of course, when physicists found energy was not being conserved in their collision experiments, a new particle with remarkable properties was quickly postulated. Initially, these were just called neutrinos. Then we needed to have three types of neutrinos. And, now, they even come with three different small rest masses. I personally would be astonished if the large-scale mass distributions observed with the Hubble telescope do not somehow have some kind of image in the space around them, and in the local speed of light. Anisotropy, speed variation in a gravitational field, changes with epoch—who knows? But now we have vastly better tools than did Michelson and Morley. Also, with a satellite mission, we can get away from the perturbing vibrations of the Earth, so it seems like an interesting idea to look around again.

Such thinking and the general technical progress have stimulated a new proposed satellite experiment, the STAR Mission. The goal is to test at the 1E–18 level frequency shifts owing to spatial anisotropy, local position invariance, gravitational potential and boost. In the first planning, the onboard optical clock of choice will use stabilization to a molecular transition in I₂ [33] or in HCCH [34].
or perhaps CO$_2$. The length etalons will be multiply redundant, with stability at the thermo-mechanical mirror motion limit. For a ULE spacer (fabricated in 1987), I now measure uniform creep below 1E–14 over the 500 s projected satellite spin period. This drift is wonderfully constant, unless there has been a recent power outage or change of temperature setpoints, or a major bump on the table. The current concept is to fit off the linear drift and examine the fit residuals for putative signals that would appear with a 250 s period.

6. Spinning off a useful secondary optical frequency standard

Planning for satellite-born experiment often leads to wildly underestimated reservoirs of error sources and other problems and, on occasion, a new road forward. Maybe I mean a thin thread to be followed. We are all aware of how valuable a resource it was to have a commercially available Cs clock/frequency standard. It worked with a broader resonance than laboratory Cs standards, and used a posteriori compensation (in abundance) to suppress unwelcome changes with temperature, pressure, atmospheric humidity, etc. Roughly speaking, one could have a long-term stability perhaps 1-order worse than the laboratory standards, but for 2-orders lower cost, size weight, need for trained staff, etc. I believe this satellite prospect may have stimulated a rough optical analogue to this situation, where saturated absorption of molecules in a high-finesse cavity can give extreme S/N, along with resonance linewidths near 1 MHz, to provide an analogous stable clock in the optical domain. It remains to be seen if the changes in the technical modulation and locking errors can be reduced to the sub-hertz domain offered by the extremely high S/N. Then a few years hence we should be able to have a secondary representation of a (new)$^2$ SI second, taken with a portable comb, it would empower frequency metrology to bring useful digits in many new areas of technology and science.

Metrology truly is the Mother of Science!

When I was employed, it was by the National Institute of Standards and Technology, and it was their support over those 44 years that made the JILA work possible. Our early laser metrology work was stimulated by JILA colleagues Peter L. Bender and James E. Faller. Of course, sometimes we would have additional support for graduate students, post-doctorates and scientific visitors via contract through the University of Colorado part of JILA. We thank the NSF, NASA, AFOSR and ONR for their help. These graduate students could do amazing things, and a number of them listened to my description of a possibly efficient research highway, then turned quietly right when I left the room and went on to do something seriously clever. Thanks to them! Especially I thank Jun Ye, whose JILA group is just speeding down the highway towards optical frequency clocks, and their applications in physics. The most essential support was by my family, especially my wife, Lindy, who put up with (most of) my research excesses over these many years. I am totally indebted to her for this friendship and support.

References

Learning from earlier SI redefinitions

11 Rosenband, T. et al. 2008 Frequency ratio of Al⁺ and Hg⁺ single-ion optical clocks; metrology at the 17th decimal place. Science 319, 1808–1812. (doi:10.1126/science.1154622)


