INSTITUTE OF PHYSICS PUBLISHING METROLOGIA
Metrologia **42** (2005) 431–441 doi:10.1088/0026-1394/42/5/014

Towards an electronic kilogram: an improved measurement of the Planck constant and electron mass

Richard L Steiner, Edwin R Williams, David B Newell and Ruimin Liu

National Institute of Standards and Technology (NIST), 100 Bureau Dr Stop 8171, Gaithersburg, MD 20899-8171, USA

E-mail: richard.steiner@nist.gov, edwin.williams@nist.gov, david.newell@nist.gov and ruimin.liu@nist.gov

Received 20 June 2005
Published 13 September 2005
Online at stacks.iop.org/Met/42/431

Abstract
The electronic kilogram project of NIST has improved the watt balance method to obtain a new determination of the Planck constant h by measuring the ratio of the SI unit of power W to the electrical realization unit W_{90}, based on the conventional values for the Josephson constant $K_{J\text{-}90}$ and von Klitzing constant $R_{K\text{-}90}$. The value $h = 6.626\,069\,01(34) \times 10^{-34}$ J s verifies the NIST result from 1998 with a lower combined relative standard uncertainty of 52 nW/W. A value for the electron mass $m_e = 9.109\,382\,14(47) \times 10^{-31}$ kg can also be obtained from this result. With uncertainties approaching the limit of those commercially applicable to mass calibrations at the level of 1 kg, an electronically-derived standard for the mass unit kilogram is closer to fruition.

1. Introduction

In 1998, we at NIST reported a value for the Planck constant h [1] as measured via the experimental technique known as the watt balance method [2]. Improving the resolution and establishing the accuracy and repeatability of measuring h are desirable in moving towards the goal of redefining the unit of mass for the International System (SI), the kilogram. Since then we have rebuilt nearly the whole experimental apparatus to achieve this improvement, finding and eliminating many old and a few new sources of error. We now report a new result that verifies our 1998 value with lower uncertainty.

2. Background

The watt balance technique measures mechanical power as defined for SI watts W against electrical power. If all measurements were in SI units, then mechanical power of force F times velocity v must equal electrical power as voltage U times current I,

$$\frac{(Fv)_{\text{mech}}}{(UI)_{\text{elec}}} \equiv 1. \qquad (1)$$

The actual procedures used illustrate that there can be a difference in these measurements and that it leads to a profound conclusion.

The watt balance apparatus and basic measurement techniques are described in greater detail elsewhere [3], so they are briefly summarized here. An induction coil (ours at room temperature) can couple in two modes to a magnetic field (0.1 T), produced in our case by a superconducting solenoid (figure 1). A trim solenoid brings the radial magnetic flux density gradient to an average $1/r$ dependence, which cancels temperature effects of first order dimensional changes in the induction coil. Running in one mode, the coil moves vertically while its velocity and the induced voltage are measured continuously and simultaneously. A quotient U/v of the measured voltage U and velocity v along the z-axis is proportional to the magnetic flux density and induction coil geometric factor. The coil voltage measurement is made against a Josephson voltage reference system [4] at near zero voltage difference with a high impedance digital voltmeter (DVM), so there is no energy loss from current heating. The coil velocity is obtained with laser interferometry as positions recorded against time.

Running in the second mode, a current I in the induction coil generates a force to statically balance the gravity force

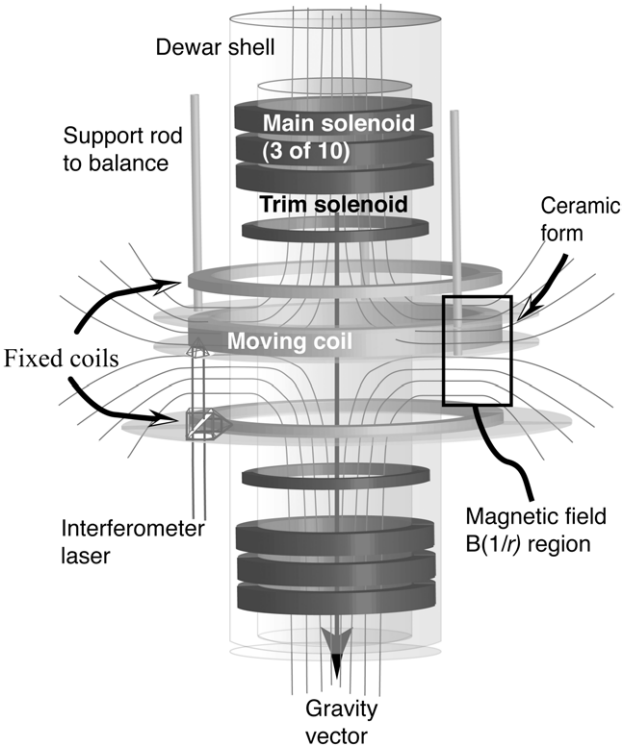

Figure 1. A schematic drawing to show the spatial relation of the magnetic flux density to the induction coils. The fixed coils are in series and in opposition with the moving coil to reduce ac pickup in velocity mode. The solenoid and the moving coil are adjusted to align electromagnetic forces with gravity to minimize horizontal force corrections.

on a reference mass ($F = mg$). A static balancing procedure is important for eliminating frictional energy losses and other hysteretic effects in the balance pivot point, a knife edge in our case. The quotient F/I is also proportional to the magnetic flux density and coil geometric factor. The vector arrow as fixed by gravity (figure 1) is a reminder that alignment of mechanical force with electromagnetic force is critical, since electrical forces or velocities not along the z-axis will not correlate with the calculated mechanical force. The local gravity acceleration g is determined by an on-site gravimeter, transferred to the height of the mass reference with a relative g-meter and estimated for real time $g(t)$ with calculated tidal corrections.

As long as the two run modes are performed quickly enough to treat the magnetic flux density as constant or linearly changing, the ratio of these two quotients is independent of the first order of the magnetic flux density and coil geometric factor. By rewriting the measurement variables in terms of mechanical power (W units SI) and electrical power (W_{90} units), we obtain the experiment's watt balance equation,

$$\frac{(Fv)_{SI}}{(UI)_{90}} = 1 + \varepsilon. \tag{2}$$

A difference factor ε is now present because electrical measurements are based on realizations that employ units V_{90}, Ω_{90} and W_{90} adopted in 1990 [5]. Voltage units V_{90} are based on the conventional value of the Josephson constant $K_{J\text{-}90}$ and the relation $U = nf/K_{J\text{-}90}$, where f is frequency and n is a quantum number. Resistance units Ω_{90} are based on the von Klitzing constant $R_{K\text{-}90}$ and the relation $Z = R_{K\text{-}90}/i$, where i is a quantum number. It is essential to make all electrical measurements consistently via these conventions. We can bring equation (2) back to an identity and show its relation with equation (1) by using a notation that separates the measurement values from the units,

$$\frac{\{Fv\}\mathrm{W}}{\{UI\}_{90}\mathrm{W}_{90}} \equiv 1. \tag{3}$$

If the measured (bracketed) value of the ratio does not equal 1, the implication is that the units relation W/W_{90} needs adjustment. W/W_{90} is the ratio of the theoretical SI values to the conventional values used for voltage and resistance measurements Z, based on the Josephson K_J and von Klitzing R_K constants, which have theoretical evaluations of $2e/h$ [6] and h/e^2 [7], respectively. Electric power U^2/Z can be written in terms of SI units, as on the left-hand side of equation (4), but is measured by convention with the defined constants, as on the right-hand side,

$$\left\{\frac{(nf/K_J)^2}{R_K/i}\right\}\mathrm{W} = \left\{\frac{(nf/K_{J\text{-}90})^2}{R_{K\text{-}90}/i}\right\}\mathrm{W}_{90}. \tag{4}$$

In this representation, n, i and f drop out, and inserting the fundamental physical constants for K_J and R_K gives

$$\frac{\mathrm{W}}{\mathrm{W}_{90}} = \frac{K_J^2 R_K}{K_{J\text{-}90}^2 R_{K\text{-}90}} = \frac{(2e/h)^2(h/e^2)}{K_{J\text{-}90}^2 R_{K\text{-}90}} = \frac{4/h}{K_{J\text{-}90}^2 R_{K\text{-}90}}. \tag{5}$$

Inserting equation (5) into equation (3) shows how the Planck constant is obtained from the measured values,

$$\frac{\{mgv\}}{\{UI\}_{90}}\frac{4}{K_{J\text{-}90}^2 R_{K\text{-}90}} = h. \tag{6}$$

From the theory of the hydrogen atom one can derive the following equation for m_e, the electron mass:

$$m_e = \frac{2R_\infty h}{c\alpha^2}, \tag{7}$$

where R_∞ is the Rydberg constant, c is the speed of light and α is the fine structure constant. Values for these and all conventional constants are obtained from CODATA 2002 [8]. Since R_∞ and α have much lower uncertainties, this measurement of h is also the most accurate measure of m_e in SI units. One can think of this experiment as indirectly counting the number of electron masses that make up one kilogram.

There are five references that underlie the SI traceability for a measurement of h. These are frequency (atomic clocks), length (lasers), voltage (Josephson effect), resistance (quantum Hall effect) and mass (artifact standard). Table 1 lists the nominal reference values used in our system. At present, the realized voltage unit is fixed, while the theoretical SI counterpart depends upon the value of h as measured in kilogram units, so any possible drift of the kilogram prototype [9] leads to greater discrepancy between W and W_{90}. One rationale for improving measurements of the Planck constant can now be established. A redefinition of the kilogram mass unit in terms of a constant h allows an all-quantum system to monitor artifact mass references, creating an electronically derived kilogram standard [10]. Improving the accuracy of this technique to about 2 parts in 10^8 is sufficient to provide long-term, commercially applicable stability to the many units that rely upon mass or force measurements.

Table 1. The nominal values for references input into the calculations, and the approximate value for various corrections measured separately from the basic data acquisition. Most of the Type B uncertainties are estimated from how well these reference values can be transferred to the experiment or how precisely the corrections can be measured.

Reference calibrations	*Magnitude*
Resistance: QHE via standard	100 Ω
Frequency: GPS and crystal	10 MHz, 5 MHz
Voltage: Josephson direct	1 V
Length: HeNe laser (iodine checked)	633 nm
Mass: Au, PtIr	1 kg
Gravity: absolute and transfer gravimeters	9.801 m s^{-2}
Corrections	*(nW/W)*
Gravity, transfer	−1430
tides	±300
pressure coefficient per hPa	±1
mass dimensions, z-position	±5
Resistor, air pressure coefficient per hPa	±4
Laser misalignment	0–10
Horizontal force (coil tilt)	±10
Horizontal force (pan displacement)	±4
Linear B field, drift per hour	<40
Crystal frequency, drift per year	−2
Abbe correction per μm offset	1

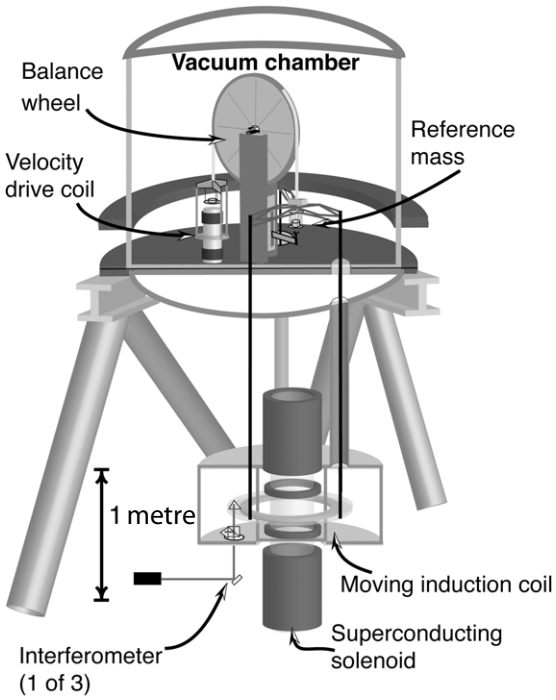

Figure 2. A schematic drawing of the complete electronic kilogram apparatus. The superconducting magnet is suspended from the upper support platform within the Dewar to keep ground vibrations common.

3. System improvements

A schematic of the complete electronic kilogram apparatus is illustrated in figure 2. Most balance hardware was reconstructed from scratch, reusing from the 1998 version of this experiment only the balance wheel, the superconducting solenoid and some optical and electronic instruments. The previous control and analysis software programs were reused and improved. A recent paper [11] describes some early failures and subsequent improvements. At that time in June 2004, most reference standards had not been fully calibrated, several tests had yet to be performed, the masses had a 200 nW/W uncertainty as an unknown experimental control and there were some suspected electrical circuit ground loops, so the uncertainty was conservatively estimated at 500 nW/W. We summarize here only the most important historical improvements and detail the most recent changes and uncertainty analysis.

The 1998 version conducted this experiment in air and achieved a relative combined uncertainty of 87 nW/W. The goal of 10 nW/W uncertainty required significant changes, and the major improvement, enclosing the coil and balance in a vacuum chamber (fibreglass to eliminate eddy currents), has made the refractive index, buoyancy corrections and air draught effects negligible. The physical apparatus is in a thermally insulated room, screened from radio interference. The tripod-base made of fibreglass composites supports both the balance and liquid helium Dewar. An aluminium optical table for the laser optics has improved laser alignment stability. Laser alignment is strongly dependent on the room temperature stability, so better room temperature control and the lagged thermal enclosure of the chamber generally hold temperatures to within 0.1 K of the set point, which also helps in maintaining the coil and balance wheel alignments. Various coil alignments are easier to set with added mechanical junctions using flexure joints and screw-adjustable translation stages. Flexures connecting the support rods to the spider and coil decouple the four degrees of coil motion of concern in alignment: two directions of pendulum swing and two of coil tilt. New testing techniques allow daily alignment of electromagnetic and optical centres and also check the coil's centre of mass line relative to that of the reference mass (a source of horizontal displacement in force mode). Some adjustments can be made while in vacuum, where an X–Y translation stage allows re-centring the induction coil should it move relative to the superconductor.

Several problems concerning the induction coil were identified when earlier-constructed coils failed to perform as expected. Vibrations are the main source of noise above 0.5 Hz in the U/v measurements, arising from sources such as seismic background, air handling equipment or liquid helium bubbling. The latest coil is stiffened against internal twisting and bending with a ceramic ring form, which has a conductive coating. The coating is etched at intervals around the coil to prevent closed-circuit eddy current loops, while segments are individually grounded to prevent static buildup. To counteract vibration effects still more, the interferometer retroreflectors are positioned halfway between a flex zero node and maximum amplitude point so that the three combined position measurements made at the coil's edge correctly represent the centre of mass motion. The dielectric properties of the wire insulation and other materials in the induction coil were also a concern. A large error arose in an early coil because the epoxy stiffener had a time-dependent dielectric response, usually referred to as dielectric relaxation. Dielectric relaxation acts as a parallel current source within the induction coil, producing an offset voltage. Constant parallel currents in both velocity and force modes will have identical offsets that cancel; a test conducted with a 10 MΩ resistor across the coil

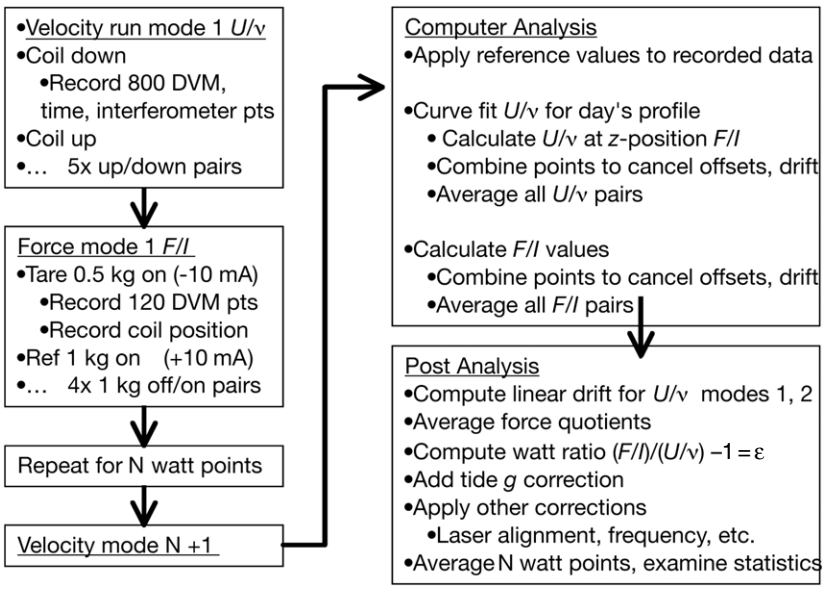

Figure 3. A flow chart of the general data acquisition (left) and data analysis (right).

in the 1998 version confirmed this. However, parallel currents from a time-dependent dielectric change will generally not cancel, becoming more severe as the relaxation time of the dielectric increases. In our velocity mode, the coil charges to 1 V about 20 s before the crucial measurement time at the centre of the translation, while in force mode the coil charges to 7 V about 60 s before the 2 min measurement begins. Modelling a correction factor is difficult, especially if assuming that the relaxation time function is an exponential. The present coil was built to minimize dielectric relaxation with a fibreglass cloth placed between each of 30 wire layers so that each cloth layer has only about 30 mV across it. This also adds space between the wire insulation itself. Most recent testing in air and after a few days in vacuum found that the dielectric relaxation decreases by a factor of 10 in vacuum. The calculated error contribution to the watt is less than 1 nW/W. In addition, we checked the calculation by modifying the velocity mode procedure for half the normal velocity. Doubling the coil translation time allows any dielectric charge a longer time to relax, thus reducing any effect by e^{-2} if the relaxation is exponential. We could see no effect within our daily scatter.

The balance wheel pivot point is an 80 mm long, sharpened wedge resting on a polished platen. This knife edge must handle $\pm 9^{\circ}$ of rotation during velocity mode. Large wheel rotations in velocity mode or small, quick ones when placing the mass on or off the balance in the force mode will cause knife edge deformations or internal stress changes. A stressed edge exhibits restoration forces with at least two recovery modes: one acting exponentially with a 2 min to 3 min recovery time and a second acting linearly over a much longer time. A cause of hysteresis in force measurements, these effects were reduced by using tungsten carbide ceramic for extra stiffness and treating both surfaces with a diamond-like carbon coating that lowers friction and adhesion [12]. Improved servo-control of the coil and mass positioning system reduced both the vertical motion of the balance to within 40 µm (about 130 mrad rotation) and the rotation time to 5 s. A procedure to erase the knife edge stress hysteresis is still necessary. After any deflection from the target angle, the wheel is put through a set pattern of rotations of decaying amplitude that especially reduces the non-linear components of the hysteresis.

4. Data analysis

Watt measurement operation, raw data recording and first pass analysis are performed automatically in software. The details have been published elsewhere, most recently in [3], so we summarize the procedure described in the background section in a flow chart (figure 3). A significant improvement in this version of the experiment arose with the superior signal/noise ratio (S/N) of the induction coil, which allowed faster identification of subtle effects, such as electromagnetic pickup from coil swinging. The signal U/v in velocity mode critically relies on simultaneous measurement of both voltage and velocity to cancel the large background 'bouncing' of the coil due to ground vibrations, mainly from nearby vibration sources or severe seismic events. In the U/v quotient, the cancellation factor is 500 for the vibration-induced voltage (few millivolt level) and velocity (few micrometre amplitude) signals at the main coil mechanical resonance frequency of 23 Hz. The S/N is about 10 000, a factor of 5 better than the 1998 system. The peak-to-peak noise of 10 µW/W equivalent is high frequency (figure 4) and cancels rapidly over the 800 readings (50 ms DVM integration per point) per trace. Each down (negative velocity) generates a -1 V signal, while an up velocity reverses all signs. All traces recorded per day are used to generate a model profile curve, calculated with an 8th order Chebychev polynomial fit. The model profile is then used to linearize each trace to obtain the U/v value at the weighing point, usually near the flat bottom of the curve where the flux density gradient has the smallest z-dependence and was adjusted for $1/r$ dependence. This profile fitting method averages all 800 points into the calculation and greatly reduces the statistical scatter. Every three consecutive traces are combined to cancel constant voltage offset and linear drift.

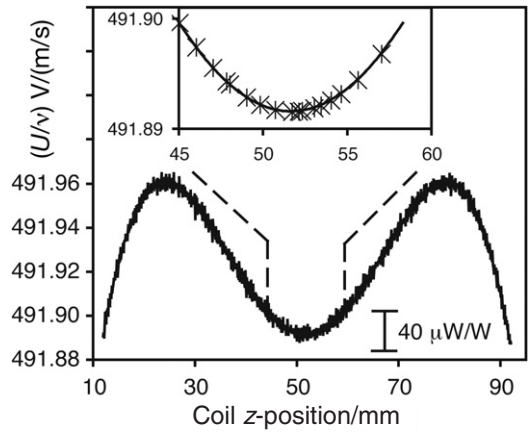

Figure 4. The profile of the magnetic flux density versus z-vertical position of the induction coil. This also shows the residual noise for one velocity pass, mainly due to residuals from voltage noise and vibrations. Vibration induces signals in both velocity and voltage, but these are coincident and rejected by a factor of about 500. The inset shows the various z-vertical test positions of the coil in the force mode. The nominal, and predominantly used, position is at the bottom of the curve to minimize non-linear effects of coil position changes, on the order of micrometres.

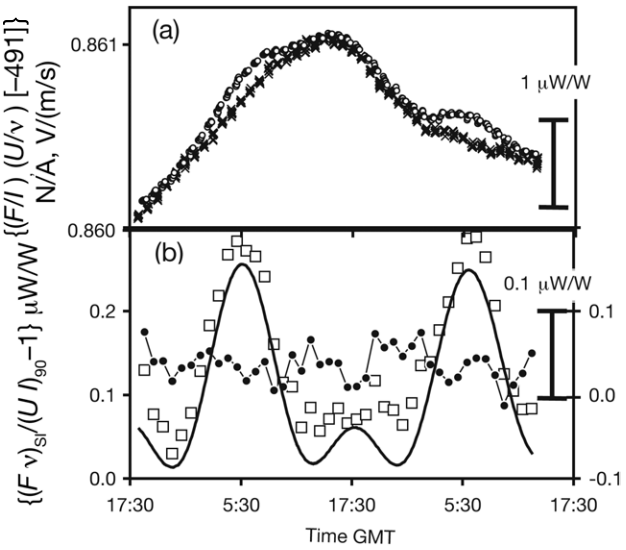

Figure 5. A two-day sample of raw data, 27 November 2004. Graph (*a*) shows how the quotient values for U/v (×) and F/I (○) vary as the magnetic flux density changes. Graph (*b*) has reduced these data to 43 watt points (□), showing how the magnetic field dependence drops out. The solid line is the calculated tidal correction to local gravity acceleration (right scale). The set of watt points (●) is shown with tides subtracted, ready for post-analysis reference corrections.

A series of five pairs of up/down coil velocity traces takes about 0.5 h and exhibits a typical sample standard deviation of about 20 nW/W equivalent.

In the force mode, a tare mass is placed on the counter mass side of the balance, where a −10 mA current-induced force pulls down to balance the coil at a desired position. This gives a convenient 1 V drop across a 100 Ω reference resistor. The current at each balance point is integrated for 2 min (120 points of 1 s DVM integration) to filter out the current source noise for a sample standard deviation of 0.1 μW/W. Improving upon the 1998 version, a faster computer servo-controls wheel angle position to a higher resolution interferometer signal but still using a retroreflector connected to the upper spider support assembly. Optimized to filter the normal coil vibrations and current source noise, the position is consistently servo-controlled to within 10 nm of the set point as averaged over 10 s or about 0.4 mrad. (The December 2004 Sumatra earthquake caused a 3 h loss of data when the slow, large seismic vibrations disrupted this servo-control but only affected one result point.) In the next step of the routine, a 1 kg mass is placed onto the same balance side as the induction coil, where a +10 mA current balances the mass. Controlling to wheel angle is essential to reduce knife edge hysteresis, since the balancing forces cause push/pull length changes of 10 μm in the long rods leading from the induction coil to the mass. These length changes are accounted for when choosing the z-value for the U/v profile. Combining three consecutive measurements again cancels voltage offsets. A series of four pairs of F/I measurements takes about 1 h. The sample standard deviation for a typical set is about 20 nW/W equivalent.

A typical watt measurement set consists of 12 to 15 repetitions of these two modes, recorded overnight, or up to 45 points over a weekend (figure 5). Upon analysis, an average of each velocity and force mode set is calculated. Since the magnetic field may drift slowly during the 1.5 h duration between U/v sets, a linear estimate is obtained from U/v averages bracketing each F/I average. The sample standard deviation of a complete set of watt values is generally within 15 nW/W to 40 nW/W. Variations between daily averages are also within that range over about 4 months of data acquisition, which included many additional tests and procedural variations looking for correlated effects, e.g. force z-axis position, integrating times or coil swing amplitude. Corrections for gravity tides, reference value updates or misalignments are added for each day's set (not all included in figure 5). An offset of up to 0.2 μW/W in all mass reference calibrations was kept unknown to us as a control on our analysis. The offset was finally revealed in January when we felt there were no more significant experimental changes to be made.

Figure 6 illustrates the improvement in the watt ratio data acquired over the last two years. These are 6023 individual watt ratio values $\{(Fv)_{SI}/(UI)_{90}-1\}$, calculated as the discrepancy ε in equation (2) in microwatts per watt. Not all corrections are included in this figure. Of note are the drastic noise reduction with the new coil and the decreasing drift of the value as the electrical grounding issues were addressed. Figure 7 shows the daily averages of the data considered for evaluation, with all corrections included. The error bars are the sample standard deviation for each set. Marked also are several procedural or environmental changes in the experiment. The magnetic field had been reversed and was set back to the previous 'normal' direction with no discernable effect. The reference mass was changed from Au to PtIr. The change at mass substitution is not considered significant, since we have a greater uncertainty for the value of the Au mass including an unmeasured vacuum surface-evaporation correction. The applied correction is estimated using the 3.7 ratio of the Au/PtIr mass surface areas to multiply the measured correction for the PtIr mass (5.5 nW/W), but it could be greater due to the rougher surface on the Au mass. Data from both masses are combined with

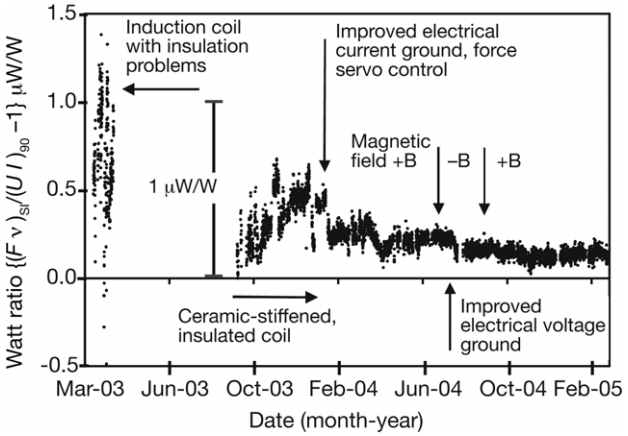

Figure 6. Over 6000 watt points showing how experiment improvements affected the data and how stable the most recent data have become. Not all corrections have been applied, especially the unknown control offset for the masses (−0.125 μW/W). Only the data taken after improving the voltage ground are in the final analysis.

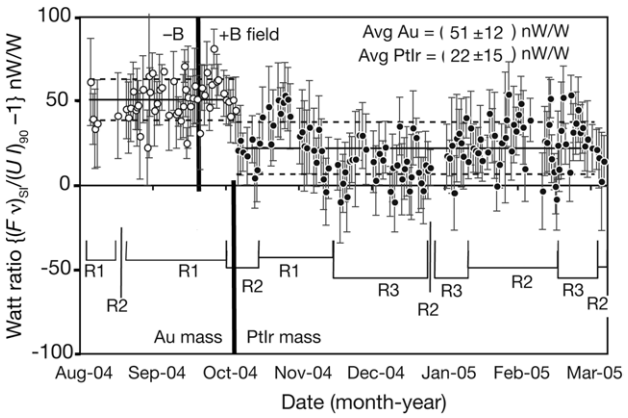

Figure 7. Analysed set of data, shown as daily averages. Error bars are the sample standard deviation of each day's points. Open circles (○) are the Au reference mass, closed circles (●) the PtIr mass. Some structure correlates to the three reference resistors used, shown as R1, R2, R3. A time of magnetic B field reversal is also indicated, with no observable effect.

weighting factors. Factors x_{Pt} and x_{Au} are calculated for both masses from the combined variances of each mass's calibration uncertainty u^2_{Pt}, u^2_{Au} and the estimated mass Type B uncertainty $u_{\mathrm{mass,\,B}}$,

$$x_{\mathrm{Pt}} = \frac{1}{u^2_{\mathrm{Pt}} + u^2_{\mathrm{mass,\,B}}}, \qquad x_{\mathrm{Au}} = \frac{1}{u^2_{\mathrm{Au}} + u^2_{\mathrm{mass,\,B}}}. \tag{8}$$

The weighting factor for each mass is normalized,

$$w_{\mathrm{Pt}} = \frac{x_{\mathrm{Pt}}}{x_{\mathrm{Pt}} + x_{\mathrm{Au}}}, \qquad w_{\mathrm{Au}} = \frac{x_{\mathrm{Au}}}{x_{\mathrm{Pt}} + x_{\mathrm{Au}}}. \tag{9}$$

Since each day's values are highly correlated, we use the number of daily means, N_{Pt}, N_{Au}. Our final watt $\mathrm{mean}_{\mathrm{watt}}$ weights the $\mathrm{mean}_{\mathrm{Pt\,Au}}$ of both masses according to the number of mass sets and the weighting factors,

$$\mathrm{mean}_{\mathrm{watt}} = \frac{\mathrm{mean}_{\mathrm{Pt}} \times w_{\mathrm{Pt}} \times N_{\mathrm{Pt}} + \mathrm{mean}_{\mathrm{Au}} \times w_{\mathrm{Au}} \times N_{\mathrm{Au}}}{w_{\mathrm{Pt}} \times N_{\mathrm{Pt}} + w_{\mathrm{Au}} \times N_{\mathrm{Au}}}. \tag{10}$$

5. Sources of uncertainty

5.1. Type A uncertainty

As mentioned above, each day's data set seems to have some internal correlation, beyond the obvious correlation to the within-set generated U/v profile, so it is better to examine the daily averaged reduced data (figure 7), which indicate some short-term structure. We suspect these patterns are correlated to short-lived variations: shifts between the reference resistors, noise affecting the velocity profile fitting routines and alignment drifts. Some corrections are only applied on a per day basis as opposed to each point. Since we acknowledge that the scatter is not 'white noise', and because of the suspected within-day dependences on average corrections, we choose to use the sample standard deviation of the daily averaged sets of data for the Type A uncertainty, rather than, for example, the standard deviation of the mean. The average value of the watt ratio $\mathrm{mean}_{\mathrm{Pt}} = (Fv)_{\mathrm{SI}}/(UI)_{90} - 1$ for all 2721 PtIr mass points taken over 174 sets is (22 ± 15) nW/W. The $\mathrm{mean}_{\mathrm{Au}}$ of 48 gold mass sets is (51 ± 12) nW/W, and the weighted $\mathrm{mean}_{\mathrm{watt}}$ as described in the paragraph above is (24 ± 15) nW/W. Even without a more sophisticated statistical analysis, this Type A uncertainty is about 10 times better than that obtained in 1998. Assignment of the Type A uncertainty is not critical because the final result is limited by Type B uncertainties.

5.2. Type B uncertainties

We classify the Type B uncertainties into four classes: basic reference, alignment, instrumental and mathematical. The basic reference uncertainties are apparent in equation (6), noting that voltage is part of the current measurement of (UI), so voltage uncertainties are squared. Though not apparent, uncertainty from a single, distributed frequency source would enter linearly but is a component in velocity (linearly), gravity and voltage measurements (both squared). Separate laser length uncertainties enter into velocity and gravity, which is included as an overall uncertainty for g. Alignment involves optical, electromagnetic and centre of mass measurements with some contributing differently in several directions, such as along the wheel surface (constant wheel radius) versus out from the surface (varying radius). Instrumental includes electrical circuitry and magnetic material property effects, most of which can be reduced with more testing and better design. Mathematical uncertainties concern procedures rather specific to this particular experiment. All units are listed as a relative uncertainty in terms of their contribution to the watt. They are listed in table 2.

5.2.1. Basic reference. The NIST Mass Measurement Group that provides mass unit traceability has moved into a new laboratory and is still re-establishing regular mass calibrations. Also, they have just started a program to measure vacuum transfers. Of the masses used, a newly constructed PtIr mass has been calibrated in air and tested for a vacuum surface-evaporation correction, both at the BIPM, but there was a 10 μg increase in a measurement performed after the vacuum test. Three NIST measurements agreed with the earlier BIPM measurement. These also indicated that the PtIr mass is

Table 2. Uncertainty budget of possible errors contributing to the watt ratio calculations. Most are estimated Type B uncertainties. Items listed as additional tests had negligible or long-period variations treatable as statistical contributions. The sample standard deviation of daily set means is used as the Type A statistical uncertainty. The relative combined standard uncertainty is the root sum square of all uncertainty components.

Basic references (error mode)	(nW/W)
Mass (force)	15
Voltage, square error (current, voltage)	2
Resistance (current)	10
Time, linear error (voltage, current, velocity)	0.5
Length (velocity)	3
Local gravity acceleration (force)	30
Alignments	
Lasers (velocity)	3
Abbe offset, optical offset × tilt (velocity)	2
Coil centres mass and EM offset, torque (force)	3
Coil tilted, translation (force)	5
Mass pan offset, translation (force)	5
Non-vertical coil motion (velocity)	2
Coil angular tilt in motion (velocity)	2
Wheel surface flatness (force)	20
Instrumental	
Electrical grounding (voltage, current)	12
Time dependent leakage resistance (voltage)	5
Laser wave front shear (velocity)	10
Mass std. magnetic susceptibility (force)	11
Superconductor flux trap hysteresis (force)	5
Superconductor-coil interaction (mode equivalence)	2
Mathematical	
Fitting order, plc change (U/v quotient)	16
Knife edge hysteresis (force)	5
Additional tests	
Interferometer timing errors	—
Non-linear drift effects	—
Polarization mixing	—
Trim current effects, asymmetric B field	—
B field reversal	—
Magnetic hysteresis ferrous metals	—
Force z-positions profile dependence	—
Type A relative uncertainty	15.6
Relative combined standard uncertainity ($k = 1$)	52

sufficiently hard, as it did not change after one month of watt balance use in 2004. The Au mass is less well characterized. It is soft and shows signs of wear, must include an air buoyancy correction and has no vacuum surface-evaporation calibration for water sorption. The combined mass uncertainty 15 nW/W equivalent is estimated from considering the calibration and vacuum surface-evaporation uncertainties, weighted relative to the amount of watt data taken for each mass.

Similarly, the resistance calibration laboratory is also being re-established but is farther along. Although measured with an uncertainty of 10 nΩ/Ω, the 100 Ω transfer standards were found to need updated characterizations of leakage resistance to case ground, pressure dependence, linear predictability and inter-lab transfer stability. Three different resistors (R1, R2, R3) were interchanged over the months. These resistors had predictable drift characteristics once transferred into the 25 °C oil bath, but also showed small shifts of a few parts in 10^8 in value upon each transfer. One resistor had a measurable pressure coefficient, so a correction was applied on a daily basis. Since there are many transfers, we treat the variations as a noise source rather than assign a large Type B uncertainty.

The gravimeter is slightly noisier than its specification of 20 nm s^{-2} [13] but has recently been repaired and is reproducing values for g equivalent to those made in 2000. Most of the assigned error is from the uncertainty in a gravity transfer measurement to a point about 5 m higher and 5 m away, where the mass reference resides. The transfer was conducted around the edge of the vacuum chamber within about 30 cm of the actual location of the mass but not with the upper vacuum chamber (about 400 kg) in operational position. Calculation indicates this is a small correction of less than 3 nW/W.

The frequency reference supplied to the Josephson voltage standard is from a GPS receiver. The listed frequency uncertainty is larger than expected for GPS, since we also use an ultrastable crystal to supply the frequency reference to the time interval analysers. This crystal source resides in an RF shielded part of the room, where we hesitated running a copper 10 MHz line from the GPS source. The crystal is checked periodically and a drift correction applied. Though affecting the Josephson system, voltage uncertainty is treated separately from frequency. The Josephson standard supplies the voltage reference directly to the watt balance instrumentation and has been verified against a second Josephson system from the volt calibration laboratory. The uncertainty is higher than usually assigned to such systems to account for thermal emfs in the leads and environmental noise that affected the system verification. The current-biased Josephson array system we use [14] allows direct connection to our coils in spite of the coils' noisy signal and so is one of the many critical improvements over the 1998 experiment.

5.2.2. Alignment. Several new routines were established to simplify alignment checks, which could then be conducted several times a week, verifying the stability of many alignment issues. Laser misalignments involve measurement of each of the three laser beams for x- and y-deviation from vertical and enter as the square of the angle. The lasers are initially aligned above the vacuum window ports and are checked external to the vacuum chamber (with a known compensation for the windows) several times a week. The external test can be verified with a more time-consuming test of misaligning each laser in turn, then running the velocity mode. The optimal setting coincides with a minimum in U/v values as a function of laser angle.

An Abbe offset error comes in as a velocity error. This error is the product of the difference between the optical centre and centre of mass times any angular coil tilt that occurs during a velocity translation. We have about 5 µrad tilt about the NS y-axis over the 40 s translation time. A high resolution method of determining a fit parameter allows us to adjust the three interferometer records to match the centre of mass, so the centres' difference is within 5 µm. This only has to be re-evaluated if changes are made to the coil.

If the induction coil shifts to the side when in force balancing mode, this creates an electromagnetic induction-related misalignment, since the coil is no longer in the same field as in velocity mode. This occurs if the mass pan is not perfectly centred over the coil's centre of mass. (There is another translation if the coil is slightly tilted.) The error enters

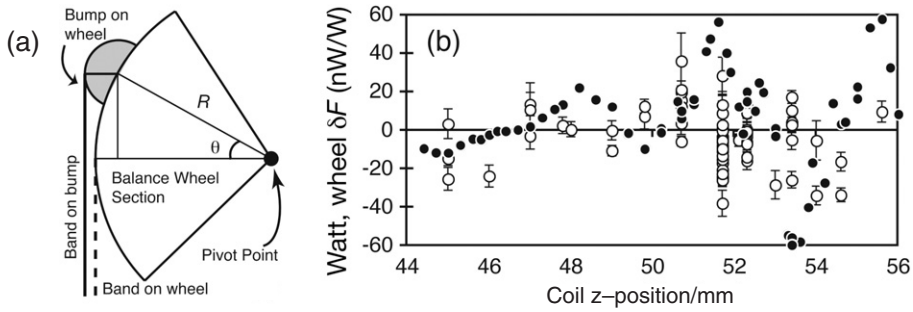

Figure 8. (*a*) A diagram showing the effect of a bump on the wheel radius R, rotating angle θ. A band hanging on the bump will result in forces sensitive to horizontal translation of the coil below. The band will displace horizontally as it sees the bump during a rotation in velocity mode. (*b*) A graph showing watt values (○) and the surface effect δP (●) as calculated from measuring the wheel surface smoothness. Note that there is no obvious correlation, especially between the 51 mm and 53 mm positions.

as the product of the side shift times the magnetic flux density gradient, which varies along that direction. Checking on this effect employs separate measurements of the displacement and the flux gradient changes along each x–y-axis. All these misalignment corrections rarely exceeded 10 nW/W. The laser alignment occasionally varied up to 80 nW/W due to room temperature excursions. The uncertainty of each alignment correction is chosen to represent the measurement resolution limits and unmeasured changes due to daily temperature variations.

The largest alignment uncertainty arises from the presumed effect of the interface between the wheel and the support band surfaces not being perfectly smooth. In reality, there are highly localized variations in the surface flatness or wire diameter. The problem of a bump on the wheel is illustrated in figure 8(*a*). A perfect wheel of radius R will have a band shown as a dashed line that is always tangential to the wheel at the $\theta = 0$ angle, so that in the force mode vertical forces on the band cause a torque about the central pivot and are insensitive to horizontal forces, while in velocity mode the velocity of the coil is purely vertical. But with a bump as shown, the band hangs down from a point that moves in an arc of varying effective radius as the wheel rotates. The balance in this mode behaves like a conventional balance where there are horizontal velocities in the velocity mode and sensitivity to horizontal forces in the force mode. Just as the NPL balance finds that these two horizontal effects cancel in the watt measurement in a conventional balance they should also cancel here. However, that assumes that the coil hanging from the bands of the balance responds ideally to the small, fast motions caused by the bumps. This is further complicated because in the velocity mode our magnetic flux density profile is fitted with an 8th order Chebychev polynomial that cannot adequately fit to small, localized changes, and thus the cancellation as in the NPL balance is less likely. In addition, pendulum motion during velocity mode also complicates any effect.

We developed a sensitive test to measure this surface effect δP with nearly nanowatt per watt equivalent resolution, essentially determining the surface smoothness by simulating a large induction coil x- or y-misalignment. The horizontal force is provided by a small side-coil mounted perpendicularly (normally used to dampen swinging). By measuring vertical force changes, then calibrating the side-coil's forces by measuring its effective U/v and F/I quotients, we can calculate δP. As expected, displacement in the y-axis direction, i.e. along the wheel radius, did have a δP component, while displacement tangential to the wheel surface had none. The map of horizontal δP error versus z-vertical position (figure 8(*b*)) reveals that for a main coil misalignment that produces a 1 µm y-axis displacement of the coil (our approximate misalignment-induced translation) there should be noticeable changes in the watt values of tens of nanowatt per watt equivalent with a granularity of a few 100 µm of z-vertical coil position. Although no position-correlated effect could be seen in the watt results recorded over various z-vertical test positions, until this is better understood, we assign a 20 nW/W uncertainty contribution for this effect.

5.2.3. Instrumental. The electrical circuitry is somewhat different from the 1998 version, and it took some time to find several instrument and electrical errors. Some suspected small ones are still difficult to determine. Our induction coil has a higher resistance of 800 Ω, so we are more sensitive now than in 1998 to voltage measurement errors from resistance leakage paths to electrical ground. An attempt to isolate the current source from computer control line grounding and reduce current noise led to worse ac coupling that induced a severe current leakage. This current source along with other instruments and all circuit paths was eventually improved to achieve leakage resistance greater than 10^{10} Ω and filter the ac coupling. The last electrical modification isolated the power supply and computer connection of the Josephson voltage system from the battery-powered control instrument connected to the watt balance circuitry. This change decreased the watt value by 60 nW/W. (This Josephson system is used directly in the watt system, whereas the previous Josephson system with voltage biasing and highly hysteretic junctions required using intermediary Zener-diode and mercury battery-based voltage sources.) Instrument modifications are now sufficient if not optimal, and the circuit is grounded in a configuration so that leakage current should not flow through critical circuit paths. After these modifications, test changes in the electrical ground relative to the Josephson array voltage source and general system ground had no measurable effect above the short-term resolution of 20 nW/W.

The magnetic susceptibilities of the Au and PtIr reference masses have not been measured, and the magnetic flux density

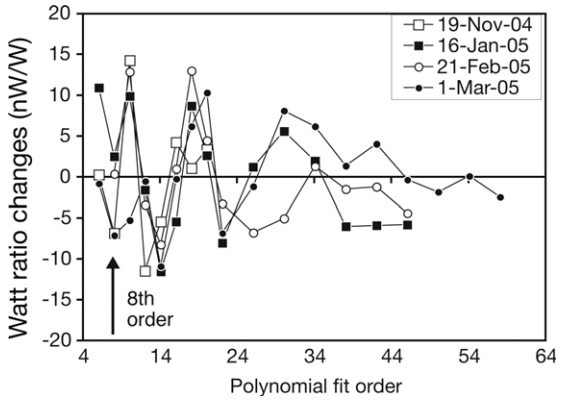

Figure 9. Effects of curve fitting the magnetic flux density profile of the velocity mode to higher order, comparing four different days. The result for this paper uses 8th order. The variation from orders 14 to 18 seems reproducible and may imply a structure within the noise of the velocity profile. Causes under further investigation include: mathematical artifacts due to noise, localized magnetic field distortions and coil horizontal motions synchronous with z-vertical, i.e. the wheel surface effects or coil swinging.

gradient at their location is not accurately known. The flux density is about 0.2 mT at the mass, so a Helmholz coil was built around the volume of the mass reference in case we need to cancel a susceptibility interaction. A test conducted with this coil to specifically look for a magnetic susceptibility effect is presently not sufficient to determine the magnitude or sign of any effect above an uncertainty of 10 nW/W.

Some uncertainty is related to the interferometry detectors. One error can occur if the beam wave front interacts with a focusing lens in the laser detector to cause a fringe shift as the beam wanders across the detector. This wave front shear would cause a velocity error if the beam moves with a systematic coil position shift. Such a coil shift occurs if our wheel rotation is off centre from the pivot point (it is not), the band slips sideways off the wheel (it does) or if there is any systematic swinging motion (there is). Our present detectors use no lenses, but the total optical system has not been well characterized, and 10 nW/W uncertainty is estimated.

5.2.4. Mathematical. A good check on several effects, including the mathematical analysis, comes from varying the force balance position along z-vertical (figure 4 inset, figure 8(*b*)). At the points farthest away from centre, the electromagnetic force differs from the preferred central region by parts in 10^5, yet the resulting watt values are within the normal scatter. In a brief test to check for any effects related to digital voltmeter timing, the normal 3 power line cycle (PLC) integration time was changed to 6 PLC. The effect was barely significant and may be related to a known sensitivity of the polynomial fitting routine to the 10 μW/W high frequency noise on the velocity profiles, as mentioned earlier. Fitting order does have some systematic effect on the watt value (figure 9) that is yet to be determined as mathematical artifact or real variations of the magnetic flux density along z-vertical. The standard deviation of a data set decreases by an order of magnitude between 5th and 6th order fitting, but does not change significantly with higher orders. This is consistent with fitting to a mathematical model of the magnetic field, derived from electromagnetic theory for solenoids. All data in this analysis are fitted to 8th order over the range 12.5 mm to 92 mm. A watt analysis routine, independently written in 2002 to check calculations in the normal data analysis, showed discrepancies of a few tens of nanowatt per watt, but velocity S/N was poor at that time due to coil problems. Simulated data generated at this same time, analysed by the normal routines, agreed to within 10 nW/W. A third test of the U/v measurement and analysis routines with synthesized frequency signals resulted in 30 parts in 10^{12} difference between the measurement analysis and calculation. Until investigation of the variations with fit order is completed, we adopt a 16 nW/W uncertainty.

As mentioned, along with the superconducting solenoid, our balance is different from others in our use of a knife edge as the balance pivot point. Other national laboratories are designing ways to entirely avoid flexing their pivot mechanisms. The hysteretic effects of our knife edge are greatly reduced from 1998, but even these smaller forces are apparent with the improved sensitivity of the new control servo-system. The highly repeatable control and erasing routine helps to significantly reduce these small effects, so the uncertainty contribution is considered to be less than 5 nW/W.

5.2.5. Additional tests. Many other tests have been performed to check for systematic errors with no effects found within the uncertainty of the test. These include laser interferometer polarization mixing, magnetic field polarity reversal, angular rotation of the coil along z-vertical and hysteretic effects of nearby ferromagnetic or diamagnetic materials.

Our interferometer uses a heterodyne technique with a dual polarized laser beam. A subtle source of error can occur if optical beam splitter misalignment allows some mixing of the polarization between the signal and reference beam [15]. Fractional measurements of the fringe are skewed by the amount of mixing to cause a periodic, non-linear error term. By making several interferometer measurements over the period corresponding to one fringe passing at normal velocity, the polarization mixing non-linearity should average out. To test for polarization mixing effects on the velocity signal, the timing was set way off from a fringe period. Increased velocity noise showed that we indeed have polarization mixing but our period timing corrects for this. In addition, our measurement timing is not fixed relative to passing fringes, so although the test measurement generated more high frequency noise, the watt value changed imperceptibly.

We do not attempt to correct for drifts that are non-linear functions of time. The worst cases of these occur (1) as an exponential function during pump-down, when outgassing in the coil rapidly changes its mass, and (2) as sudden-onset transitions during temperature control failures, when alignments suffer. Both of these also cause a temperature gradient across the wheel, which rapidly alters its dimensions. These problematic data sets are not included in the analysis. They do indicate that a linear approximation in the analysis handles the normal, small non-linear drifts adequately but also highlights extremely non-linear drifts or changes with an increased noise figure. Drift of the magnetic field can have a slowly varying non-linear component, but the direction varies

from day to day, so we assume it adds random noise. Tests for non-linear contributions due to nearby ferromagnetic materials showed that the effects cancel, as long as they do not move. But there is also a question about the effect of the induction coil on the superconducting solenoid, so the following test was developed.

This important new test that answers the question 'Do we know the field is the same for the velocity mode and the force mode?' is as follows: current I_f is injected into the pair of fixed induction coils to produce a flux change within and a force on the superconductor that is similar to that produced by the moving induction coil in the force mode. With the magnetic field on, we measure the small force between the fixed coils and moving coil at various z-vertical positions, and for convenience find the z-position where the force between the fixed and moving coils is zero. Near this zero position we set current in the moving coil to $+I_m$ and $-I_m$ and measure the force change for $\pm I_f$ in the fixed coil. The force magnitude should be the same for either polarity I_m, but we detect a 300 nW/W equivalent difference that we credit to two effects: the change in the dimensions in the superconductor due to large magnetic-coupling forces and also the diamagnetism of the superconductor. Then we repeat this procedure with no current in the superconductor while it is still cold. This yields a difference between $\pm I_m$ force tests of 50 nW/W equivalent, which we believe is due to diamagnetic forces only. The averages of the $\pm I_m$ tests, with and without superconducting current, differ by only 2 nW/W equivalent. Thus, we find the superconducting magnetic field does change with current in either the fixed or moving coil (about 2 parts in 10^7), but it is linear with moving-coil current. Therefore, the answer to the question posed at the beginning of this paragraph is that the average of the magnetic field over a *balanced* +10 mA and −10 mA current in the induction coil is within a few parts in 10^9 of the field when there are no current-induced forces, i.e. the average force mode field is *the same* as the velocity mode field. It also implies that a symmetric current-reversing procedure is required. As mentioned above, a probable reason for most of the small change in field is that, in reaction to the 10 N force on the induction coil, the superconducting solenoid responds with a small dimensional change (a permanent magnet will also likely have this) or a small shift in the supercurrent distribution within the superconducting wire. This was expected and now we have proved it to be linear, so no corrections are required. The authors are indebted to Joshua Schwarz for proposing this test.

6. Summary and outlook

From our results for the watt ratio $(Fv)_{SI}/(UI)_{90} - 1 = W_{90}/W - 1 = (24 \pm 52)$ nW/W and from equation (6), we derive a value for the Planck constant $h = 6.626\,069\,01(34) \times 10^{-34}$ J s. Similarly from equation (7), the electron mass $m_e = 9.109\,382\,14(47) \times 10^{-31}$ kg. The superior alignment and S/N, along with the reduced dependence upon air and temperature effects of this new experiment, revealed no significant discrepancy with the NIST 1998 version of the experiment [1]. For one error source that we did not consider in 1998, the dielectric relaxation in the induction coil, we measured the old coil. The result suggests a reduction of the 1998 value by about 25 nW/W, but this difference is still within even our present uncertainty. These results also agree with the NPL 1988 result [16], but there is a large 1078 nW/W discrepancy with the Avogadro-derived value of 2004 [17] (figure 10).

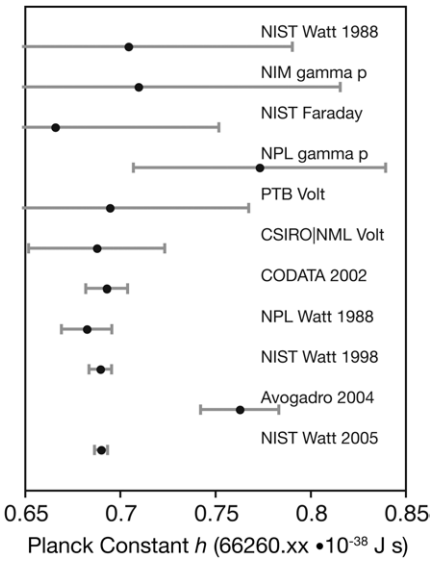

Figure 10. A comparison of this new Planck value and previously reported values. Note the large discrepancy between the watt balance and Avogadro-derived values.

We hope to reduce the combined relative uncertainty by about a factor of 2 by the end of 2006. The main focus will be to reduce the largest Type B uncertainties related to gravity, mass, wheel radius effect and most of the instrumental uncertainties. It should be possible to change the experiment design so that some effects, such as resistance standard transfers, z-vertical alignment dependences and polynomial fit order, can be identified as contributing to the Type A uncertainty. Many points of the present data set include results taken over a variety of test parameters, so extended data acquisition under constant run parameters may allow more sophisticated statistical analyses to determine the type of noise and thus a reduction in the Type A uncertainty. Our goal of better than 20 nW/W uncertainty and reproducibility seems achievable in the near future.

Acknowledgments

We are grateful to R Elmquist, J Sims, C Burroughs, Y Tang, J Keller, Z Jabbour of NIST and D Winester of NOAA, Boulder, CO, for their assistance in calibrations vital to this experiment. The contributions to the construction and implementation of the experiment by A Picard of BIPM, P Gournay of BNM, Paris, France, J Schwarz of NOAA, Boulder, CO and P T Olsen (ret.) are also gratefully acknowledged.

References

[1] Williams E R, Steiner R L, Newell D B and Olsen P T 1998 Accurate measurement of the Planck constant *Phys. Rev. Lett.* **81** 2404–7

[2] Kibble B P 1976 A measurement of the gyromagnetic ratio of the proton by the strong field method *Atomic Masses and Fundamental Constants* vol 5, ed J H Sanders and A H Wapstra (New York: Plenum) pp 545–51
[3] Steiner R, Newell D and Williams E 2005 Details of the 1998 watt balance experiment determining the Planck constant *J. Res. Natl. Inst. Stand. Technol.* **110** 1–26
[4] Taylor B N, Parker W H, Langenberg D N and Denenstein A 1967 On the use of the ac Josephson effect to maintain standards of electromotive force *Metrologia* **3** 89–98
[5] Taylor B N and Witt T J 1989 New international electrical reference standards based on the Josephson and quantum Hall effects *Metrologia* **26** 47–62
[6] Josephson B D 1962 Possible new effects in superconductive tunneling *Phys. Lett.* **1** 251–3
[7] von Klitzing K, Dorda G and Pepper M 1980 New method for high accuracy determination of the fine-structure constant based on quantized Hall resistance *Phys. Rev. Lett.* **45** 494–7
[8] Mohr P J and Taylor B N 2005 CODATA recommended values of the fundamental physical constants: 2002 *Rev. Mod. Phys.* **77** 1–107
[9] Quinn T 1995 Base units of the Systeme International d'Unités, their accuracy, dissemination and international traceability *Metrologia* **31** 515–27
[10] Mills I M, Mohr P J, Quinn T J, Taylor B N and Williams E R 2005 Redefinition of the kilogram: a decision whose time has come *Metrologia* **42** 71–80
[11] Steiner R L, Newell D B, Williams E R, Liu R and Gournay P 2005 The NIST project for the electronic realization of the kilogram *IEEE Trans. Instrum. Meas.* **54** 846–9
[12] Schwarz J P, Liu R, Newell D B, Steiner R L, Williams E R and Smith D 2001 Hysteresis and related error mechanisms in the NIST watt balance experiment *J. Res. Natl. Inst. Stand. Technol.* **106** 627–40
[13] Niebauer T M, Sasagawa G S, Faller J E, Hilt R and Klopping F 1995 A new generation of absolute gravimeters *Metrologia* **32** 159–80
[14] Benz S P, Hamilton C A, Burroughs C J, Harvey T E and Christian L A 1997 Stable 1 volt programmable voltage standard *Appl. Phys. Lett.* **71** 1866–8
[15] Wu C and Deslattes R 1998 Heterodyne interferometer with subatomic periodic nonlinearity *Appl. Opt.* **37** 6696–700
[16] Kibble B P, Robinson I A and Belliss J H 1990 A realization of the SI watt by the NPL moving-coil balance *Metrologia* **27** 173–92
[17] Fujii K *et al* 2005 Present state of the Avogadro constant determination from silicon crystals with natural isotopic compositions *IEEE Trans. Instrum. Meas.* **54** 854–9