IOP Publishing *Metrologia* **63** (2026) 025012 https://doi.org/10.1088/1681-7575/ae570f

Metrologia

RECEIVED
7 January 2026

REVISED
24 March 2026

ACCEPTED FOR PUBLICATION
25 March 2026

PUBLISHED
16 April 2026

PAPER

Redetermination of the gravitational constant with the BIPM torsion balance at NIST

Stephan Schlamminger[1,*], **Leon Chao**[1], **Vincent Lee**[1], **Craig Shakarji**[1], **Antonio Possolo**[1], **David Newell**[1], **Julian Stirling**[2], **Robert Cochrane**[3] and **Clive Speake**[4]

[1] National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD, United States of America
[2] Humanitarian Technology Trust, Bath, United Kingdom
[3] Independent researcher, Stirling, United Kingdom
[4] School of Physics and Astronomy, University of Birmingham, Birmingham, United Kingdom
[*] Author to whom any correspondence should be addressed.

E-mail: stephan.schlamminger@nist.gov

Keywords: gravitational constant, torsion balance, precision measurement, thermophoresis, small torques

Abstract

We report the first replication of a high-precision measurement of the gravitational constant, G. The experiment employed the torsion balance originally designed and constructed at the International Bureau of Weights and Measures (BIPM) approximately three decades ago. Using the same apparatus and geometry, with several modifications documented in this work, we determined $G = (6.67387 \pm 0.00038) \times 10^{-11}\,\mathrm{m}^3\,\mathrm{kg}^{-1}\,\mathrm{s}^{-2}$, corresponding to a relative standard uncertainty of 5.7×10^{-5}. The result is lower by 2.5×10^{-4} relative to the BIPM determination. This replication provides an independent verification of one of the most precise torsion-balance determinations of G and contributes to assessing the reproducibility limits of current experimental techniques in measurements of the gravitational constant.

List of Acronyms

BIPM	International Bureau of Weights and Measures.
CMM	Coordinate measuring machine.
CODATA	Committee on Data of the International Science Council.
CTE	Coefficient of thermal expansion.
MC	Mounting circle.
MCR	Mounting circle radius.
MPE	Maximum permissible error.
NIST	National Institute of Standards and Technology.
PTB	Physikalisch-Technische Bundesanstalt.
RMS	Root mean square.
SM	Source mass.
TM	Test mass.

List of Symbols

α	phase difference between electrode voltages V_1 and V_2
α_{fid}	direction of SM density gradient relative to a fiducial mark on the mass
α_{lab}	direction of SM density gradient in lab
α_{T}	coefficient of thermal expansion (CTE)
β	velocity dependent damping constant
$\boldsymbol{\Sigma}, \mathbf{R}$	covariance and correlation matrices
$\Delta\phi$	angular difference $\phi_{\mathrm{s}} - \phi_{\mathrm{t}}$
Δz	vertical difference of assemblies $z_{\mathrm{s}} - z_{\mathrm{t}}$
$\Delta\rho$	gradient of source mass density
η	dimensionless ratio $R_{\mathrm{t}}/R_{\mathrm{s}}$
$\Gamma(\Delta\phi)$	dimensionless mass integration factor
Γ_{s}	value of Γ using R_{s}
κ	torsion constant
κ_{e}	elastic part of the torsion constant
κ_{g}	gravitational part of the torsion constant
κ_{i}	imaginary part of the torsion constant
κ_{r}	real part of the torsion constant
λ	thermal conductivity of residual gas

$\lambda_j, \tau_j, \mu, \varepsilon_j$	Parameters for the Bayesian hierarchical model (configuration effect, dark uncertainty, consensus value, error)
$\mathbf{C}_{\mathrm{cap}}$	symmetric capacitance matrix
$\mathbf{V}$	vector of electrode potentials $\mathbf{V} = (V_1, V_2, V_3)^{\mathsf{T}}$
ω	angular frequency
ω_0	angular resonance frequency of the pendulum
ϕ_{max}	ϕ_{s}, where $N(\Delta\phi)$ is maximal, also $\Gamma = \Gamma_{\mathrm{max}}$
ϕ_{min}	ϕ_{s}, where $N(\Delta\phi)$ is minimal, also $\Gamma = -\Gamma_{\mathrm{max}}$
p_{v}	residual gas pressure in the vacuum chamber
ρ	source mass density average value
τ	time constant
τ_0	shortest time for anelastic relaxation
τ_∞	longest time for anelastic relaxation
θ	orientation angle of the source masses (clocking)
θ_0	phase of the source mass density inhomogeneity
ξ	modulus defect
b_{st}	width of the torsion strip
$c_x(y)$	relative sensitivity of y w.r.t. x, i.e. $\frac{x}{y}\frac{\partial y}{\partial x}$
$C_{\mathrm{c},ij}$	cross capacitance between electrode i and j.
C_{ij}	mutual capacitance between electrodes i and j.
d_m	diameter of a gas molecule
E_{el}	electro static energy
E_{pot}	potential energy of the torsion pendulum
f	number of degrees of freedom
F_0	shear modulus of the torsion strip
F_x	thermophoretic force in x direction
G	gravitational constant
g	local acceleration
I_{disk}	moment of inertia of the disk
I_{in}	moment of inertia, brass balls in
I_{out}	moment of inertia, brass balls out
I_{t}	moment of inertia of test masses about fiber
k	Boltzmann constant
k_{ij}	capacitance gradient, $k_{ij} = \frac{\mathrm{d}C_{\mathrm{c},ij}}{\mathrm{d}\phi_{\mathrm{t}}}$
l	mean free path
l_0	mean free path in an infinite volume
l_{st}	length of the torsion strip
l_{typ}	typical dimension of a vacuum can
m_{disk}	mass of the disk
m_{g}	molecular mass of residual gas
$N(\Delta\phi)$	gravitational torque on the pendulum bob
N_{ext}	external torque acting on the pendulum
P_l, P^1_{4l}	Legendre polynomials and associated Legendre polynomials
p_Q	product p_{v} with Q_{gas}
Q	observed quality factor of the pendulum
Q_{gas}	quality factor due to gas damping
Q_n	robust estimator of standard deviation
Q_{max}	quality factor achievable for $p_{\mathrm{v}} = 0$
r_{eff}	effective lever arm of the torsion balance
r_{sph}	radius of a sphere suspended in vacuum
T	temperature
T_{in}	period of the pendulum, brass balls in
T_{out}	period of the pendulum, brass balls out
t_{st}	thickness of the torsion strip
$u(x)$	$(k=1)$ uncertainty of quantity x
V_i	potential on electrode i
$V_{1,\mathrm{rms}}$	RMS voltage on 1, $V_{1,\mathrm{rms}} = V_{\circ,\mathrm{rms}} + V_{\Delta,\mathrm{rms}}$
$V_{2,\mathrm{rms}}$	RMS voltage on 2, $V_{2,\mathrm{rms}} = V_{\circ,\mathrm{rms}} - V_{\Delta,\mathrm{rms}}$
$V_{\circ,\mathrm{rms}}$	common-mode voltage
$V_{\Delta,\mathrm{rms}}$	differential voltage
W	weight suspended from the torsion strip
a_{t}	bore radius of test mass
b_{s}, b_{t}	outer radii of source and test masses
Γ_{max}	maximum value of Γ
G_{f}	G obtained with the free-deflection method
G_{s}	G obtained with the servo method
h_{s}, h_{t}	heights of source and test masses
m_{s}, m_{t}	mass of one source and one test mass
ϕ_{s}	azimuthal angle of the source mass assembly
ϕ_{t}	azimuthal angle of the test mass assembly
R_{s}, R_{t}	mounting-circle radius of SM and TM
z_{s}	vertical location of source mass center of mass
z_{t}	vertical location of test mass center of mass

1. Motivation and background

The Newtonian gravitational constant, G, sets the strength of the gravitational interaction and is one of the fundamental constants that define our Universe. Despite this central role, G is known with unusually large uncertainty. The large uncertainty in G stems from mutually inconsistency of the measurement results [1–18] obtained over the past five decades (figure 1). This persistent scatter points to two possibilities: new physics or unrecognized instrumental effects. The first hypothesis, unexplained physics, remains remote. The second, more plausible one, is that subtle instrumental systematics underlie the divergence among results. While progress on the first possibility could deepen our understanding of gravity and its remaining mysteries, clarifying the second would advance the development of precision instruments used to measure small forces, torques, and other subtle signals relevant to fundamental physics. In any case, improving our understanding of the challenges involved in measuring G is a worthwhile endeavor.

One way to improve knowledge of G is to design new, more precise experiments. This strategy was pursued through an Ideas Lab organized by the National Science Foundation and hosted at the NIST in 2016 [20]. Yet adding more measurements does little good: the earlier, inconsistent results still dominate the weighted mean, and additional data alone cannot resolve the overall tension among the existing measurements. A better strategy is to re-examine outliers [21]. This idea gained traction during a 2014 NIST workshop [22]. A photograph of the workshop participants is provided in figure 2.

At the meeting, Terry Quinn, the principal investigator of the experiment at the BIPM, proposed to send his torsion balance to NIST so that an independent team could attempt a measurement of G using the same instrument.

The suggestion was apt: the BIPM results (figure 1) were among the most divergent from the value that CODATA recommended most recently. Replicating the experiment at NIST offered a chance to probe the design itself and to possibly unveil any hidden systematics.

The determination reported here is not entirely independent, as C. C. Speake contributed to both the BIPM and NIST efforts. He visited NIST several times to help set up and discuss the apparatus, and remained actively involved through regular remote meetings over nearly a decade.

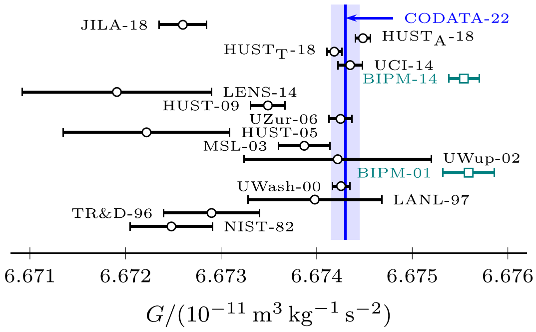

Figure 1. Almost half a century of measurements of the gravitational constant G with ($k = 1$) uncertainty bars. The solid vertical line and blue shaded area indicate the 2022 Committee on Data of the International Science Council (CODATA) [19] recommended value and its stated uncertainty. Teal squares denote measurements made with the International Bureau of Weights and Measures (BIPM) torsion balance. Labels follow the CODATA convention and refer to the following references: NIST-82: [1], TR&D-96: [2], LANL-97: [6], UWash-00: [3], BIPM-01: [4], UWup-02: [7], MSL-03: [8], HUST-05: [9, 10], UZur-06: [11], HUST-09: [12, 13], LENS-14: [16], BIPM-14: [5], UCI-14: [17], HUST$_T$-18: [18], HUST$_A$-18: [18], JILA-18: [14, 15].

2. History of the BIPM experiment

The BIPM experiment emerged as the culmination of three distinct developments in the late 1990s. (1) Researchers at the BIPM had extensively studied [23–26] the loss mechanisms in copper–beryllium (Cu–Be) flexure strips, originally to develop an improved mass comparator [27, 28] for disseminating the kilogram [29] and for fundamental-physics measurements [30]. (2) Kuroda's analysis [31] of anelasticity in torsion fibers cast doubt on the accuracy of several published G measurements that relied on the time-of-swing method, the most precise technique available at the time. (3) Researchers at the PTB in the former West Germany published [32–34] a value of G that was more than 40 standard deviations above the then-accepted CODATA value [35], spurring a wave of new measurements of the gravitational constant—though the PTB result was later retracted [36] following modeling by Speake [37].

Out of the combination of these factors grew a proposal [38] for a novel torsion balance to measure the Newtonian gravitational constant that utilized a heavily loaded Cu–Be strip. As discussed in the proposal, this design marked a departure from C. V. Boys' approach [39], which favored a small torsion bob suspended by a very fine wire with circular cross section to maximize the signal-to-noise ratio.

Instead, Drs. Quinn, Speake, and Davis advocated the use of a rectangular torsion strip, taking advantage of an effect now known in the literature as dissipation dilution [40]. For a torsion strip, mechanical dissipation (the conversion of mechanical energy resulting from twisting or oscillation into thermal energy) is relatively smaller than for a round fiber due to the contribution of an additional, lossless gravitational restoring torque.

Since thermal noise is proportional to the lossy part of the spring constant, increasing the TM, and with it the gravitational torque can improve the

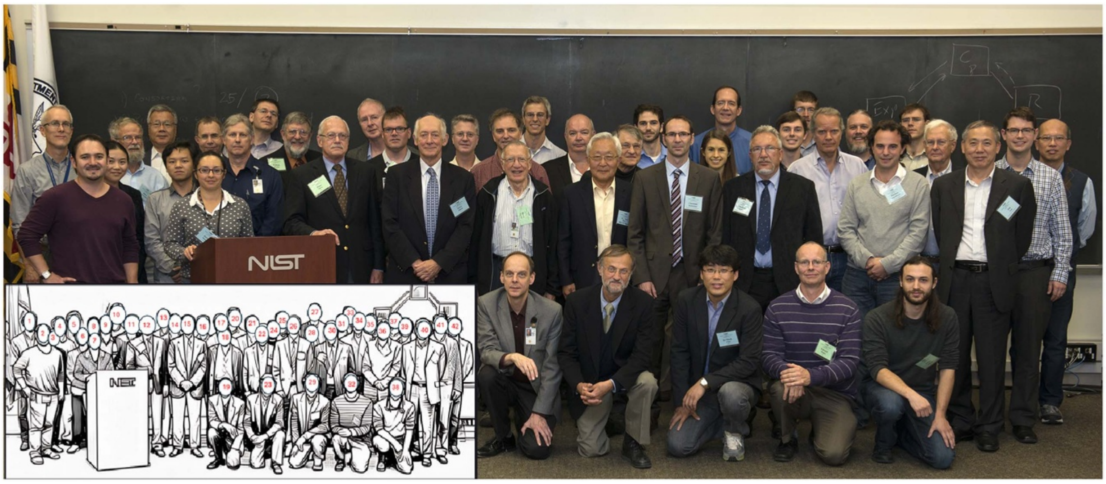

Figure 2. Group photo by Jennifer Lee (NIST) of some of the attendees of the workshop on the gravitational constant held at NIST on November 7, 2014. 1- Jon Pratt 2- Felipe Guzman 3- Qinglan Wang 4- Riley Newman 5- Kazuaki Kuroda 6- Xunnong Xu 7- Darine Haddad 8- John Gillespie 9- Richard Steiner 10- Seth Aubin 11- Bill Phillips 12- Peter Mohr 13- Ian Robinson 14- Stephan Schlamminger 15- Terry Quinn 16- David Newell 17- Jens Gundlach 18- Barry Taylor 19- Carl Williams 20- Joshua Bienfang 21- Barry Wood 22- Ho Jung Paik 23- Andrea De Marchi 24- Klaus von Klitzing 26- Christian Rothleitner 27- Richard Allen 28- Christine Donnelly 29- In-Mook Choi 30- Vladimir Nesterov 31- Chris Overstreet 32- Edward Hagley 33- Ira Thorpe 34- Tom LeBrun 35- Gordon Shaw 36- Markus Aspelmeyer 37- Charles Hagedorn 38- Julian Stirling 39- James Faller 40- Jun Luo 41- John Melcher 42- Joseph Tan. The authors have confirmed that any identifiable participants in this study have given their consent for publication.

signal-to-noise ratio. A rectangular torsion strip can support a greater load for the same absolute level of internal dissipation than a round torsion fiber. Consequently, the gravitational torque to be measured by this torsion balance is substantial, approximately 30 nN m, roughly 10^4 times larger than in typical torsion balance experiments.

Another distinctive feature, already present in the first prototype, was the fourfold symmetry of the TMs and SMs. This hexadecapole gravitational configuration offers a key advantage: the gravitational coupling from the SMs decreases with distance as ${R_s}^{-5}$, where R_s is the radius of the MC of the SMs. This steep falloff suppresses unwanted couplings from external masses in the laboratory, which diminish just as rapidly with distance. Because the torque scales as ${R_s}^{-5}$, even small errors in the radius propagate strongly, requiring dimensional metrology to be carried out with a CMM.

The first prototype torsion balance incorporating the features described above was already capable of detecting gravitational signals in 1997 [38]. By 1998, at the bicentennial celebration of Cavendish's measurement, a complete science version, later termed *Mark I*, was introduced to the community [41].

At the bicentennial meeting, the audience heard about another distinctive feature of the BIPM experiment. The gravitational constant was measured in two ways: the free-deflection method, also known as the Cavendish method, and the electrostatic servo method. This dual-methode capability was intended not only to produce a more robust determination of G but also to shed light on the persistent scatter in the global dataset of gravitational constant measurements.

Just three years after its introduction, a precision result obtained with *Mark I* was published [4] with a relative uncertainty of 4.1×10^{-5}. The final value combined results from both measurement methods, which showed good agreement. Notably, the G measurement was 2×10^{-4} higher than the best measurement at the time [3].

Mark II was a rebuilt version with a sturdier frame, a direct-view autocollimator (eliminating angle multiplication), a more precise CMM, and SMs shortened by 2 mm to recut the damaged kinematic mounts. Different TMs were used in *Mark II*, and the electrostatic shielding was improved by concealing all insulators beneath an electrically conducting false floor.

Most of the data collected with *Mark II* were obtained in late 2007 [42], but the final result did not appear until 2013 [43], followed by an erratum in 2014 [44]. With a relative uncertainty of 2.4×10^{-5}, *Mark II* achieved nearly half the uncertainty of *Mark I*. As before, both measurement methods were combined to obtain the final result. While the methods remained consistent within their correlated uncertainties, the agreement was less striking than in 2001. Most importantly, however, the results from *Mark I* and *Mark II* are in good agreement.

A Royal Society discussion meeting titled *The Newtonian Constant of Gravitation* was held at Chicheley Hall in February 2014 [45]. It led to a special issue of Philosophical Transactions A containing the full *Mark II* description [5].

The apparatus was shipped to NIST in 2016, where its new implementation is described next. This work represents the first independent replication of a G experiment using the original hardware but mostly new personnel, environment, and data analysis pipeline. The goal was not to improve the world average of G, but to identify and quantify the hidden systematic effects that limit present-day precision.

3. Measurement technique

Having reviewed the historical background, we now describe how the experiment operates. We first define the geometry and torque (equation (1)), then derive the two operating modes: free deflection (equation (10)) and electrostatic servo (equation (20)) Finally, we compare the sensitivities of their parameters (table 1).

3.1. The gravitational torque

Figure 3 shows the mass geometry. Four cylindrical copper TMs, each of mass m_t, are mounted on an aluminum disk and spaced 90° apart on a MC of radius R_t. The torsion disk is suspended by a torsion strip with spring constant κ. The torsion strip hangs from a two-dimensional gimbal system, which decouples the top clamp from ground tilts. Eddy current dampers are then used to dampen oscillations.

The pendulum's total moment of inertia is $I = I_\mathrm{t} + I_\mathrm{disk}$, where I_t is due to the four TMs and I_disk to the torsion disk. The latter includes the disk itself, mirror tower, bottom clamp for the torsion strip, eight brass screws, and small leveling balls; we refer to these collectively as the *disk*. All components are housed in a grounded vacuum can (not shown), which serves as an electrostatic shield between the pendulum and the external SMs to prevent electrostatic interactions. Inside the vacuum chamber, and also shown in the figure, are eight copper rods. Four rods are electrically connected to form electrode 1, and the remaining four form electrode 2. Applying a potential between the TMs (electrode 3) and either electrode 1 or 2 generates a counterclockwise or clockwise torque on the pendulum, respectively.

Outside the vacuum envelope, four SMs, each of mass m_s, are spaced 90° apart on a MC of radius R_s. These are mounted on a carousel that can be rotated using a stepper motor. Numerical values of key parameters appear in table 2.

Two azimuthal angles describe the geometry: ϕ_s for the SM assembly and ϕ_t for the pendulum. The gravitational interaction between the SM and TM assemblies depends on their relative angular displacement, defined as $\Delta\phi = \phi_\mathrm{s} - \phi_\mathrm{t}$ (see section 6):

$$N(\Delta\phi) = G\frac{8 m_\mathrm{s} m_\mathrm{t}}{R_\mathrm{s}}\Gamma(\Delta\phi), \tag{1}$$

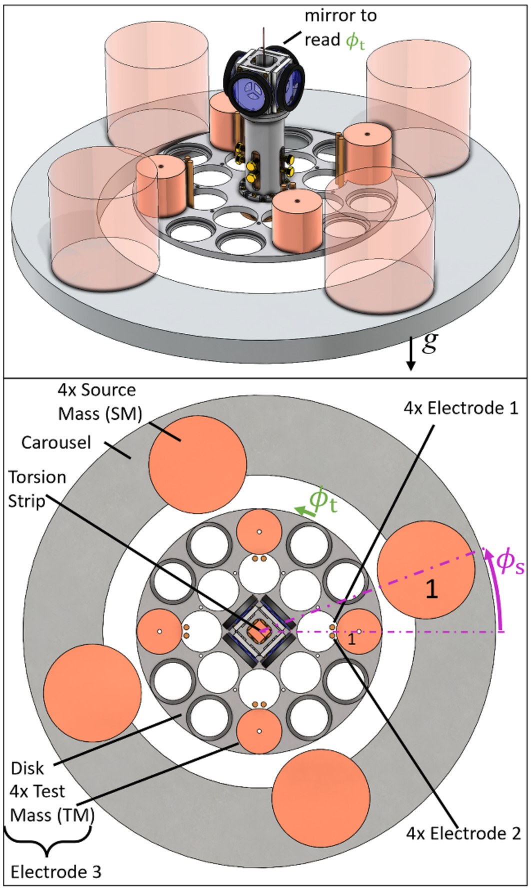

Figure 3. Layout of the BIPM experiment. The torsion disk, including the TMs, is housed inside a vacuum can (omitted for clarity to reveal the internal components). The azimuthal angle of the torsion disk is denoted by ϕ_t (shown here as zero), and that of the SM assembly by ϕ_s. As indicated in the figure, all angles are measured positively in the counterclockwise direction, as viewed from above.

where $\Gamma(\Delta\phi)$ is a dimensionless mass-integration factor with maximum ≈ 0.485 for this geometry.

To maximize signal, the SM assembly is modulated between two azimuths, ϕ_max and ϕ_min, separated by 37.67°, corresponding to the extrema of torque on the TM assembly. The positions of the extrema are determined by finely dithering ϕ_s and locating zeros of $\partial N/\partial\phi_\mathrm{s}$:

$$\frac{\partial N(\phi_\mathrm{s} - \phi_\mathrm{t})}{\partial\phi_\mathrm{s}} = 0. \tag{2}$$

We define

$$\Gamma_\mathrm{max} = \Gamma(\phi_\mathrm{max} - \phi_\mathrm{t}) \tag{3}$$

and

$$\Gamma_\mathrm{min} = \Gamma(\phi_\mathrm{min} - \phi_\mathrm{t}) \approx -\Gamma_\mathrm{max}. \tag{4}$$

Combining the torque definition and geometry factor, the gravitational constant follows directly as

$$\boxed{G = \frac{\Delta N_\mathrm{max} R_\mathrm{s}}{8\, m_\mathrm{s} m_\mathrm{t}(\Gamma_\mathrm{max} - \Gamma_\mathrm{min})} \approx \frac{\Delta N_\mathrm{max} R_\mathrm{s}}{16\, m_\mathrm{s} m_\mathrm{t}\Gamma_\mathrm{max}}} \tag{5}$$

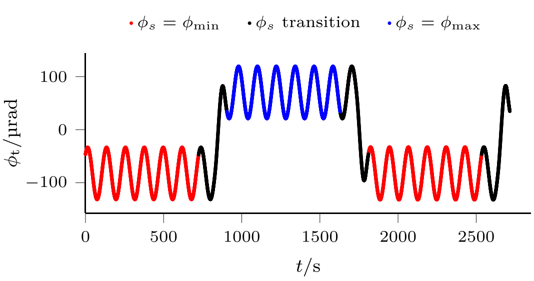

Figure 4. Autocollimator trace of the torsion angle ϕ_t recorded in the free-deflection method. The SMs begin at $\phi_{\min}$ (red points) and data are recorded for six pendulum periods. They are then moved to $\phi_{\max}$ (blue points) for another six periods. The motion between positions is controlled to maintain a constant oscillation amplitude of 48 µrad (see [46]). Data taken during the transition (black points) are excluded from the analysis.

The second form uses $\Gamma_{\min} \approx -\Gamma_{\max}$, a good approximation for the symmetry of this geometry.

The typical values for $\Delta N_{\max}$ are 31.2 nN m and 14.0 nN m for the experiments using copper and sapphire SMs, respectively. Precise values are listed in table 15. To measure these torques with a relative uncertainty of 50×10^{-6}, the absolute uncertainty must be 1.6 pN m and 0.7 pN m, respectively. For comparison, recall that the signal in one of the most precise measurements of G is only $\approx$ 12 pN m [3].

We measure $\Delta N_{\max}$ by two methods: free deflection and electrostatic servo (below).

3.2. The free deflection method

For a free torsional pendulum, modulating the gravitational torque $\Delta N_{\max}$ produces a peak-to-peak equilibrium shift

$$\Delta\phi_t = \frac{\Delta N_{\max}}{\frac{\partial^2 E_{\mathrm{pot}}}{\partial {\phi_t}^2}}, \tag{6}$$

where the second derivative in the denominator is explained below.

A typical segment is shown in figure 4. The autocollimator records ϕ_t relative to an arbitrary reference. The SMs start at $\phi_{s,\max}$ for six periods, then move to $\phi_{s,\min}$ for another six. Motion between positions is controlled to maintain a constant oscillation amplitude of 48 µrad (see [46]). Although data are taken continuously, only intervals with stationary SM positions are used in the analysis. This back-and-forth modulation repeats over several days.

For each SM position, we fit a sinusoidal function directly to the time series of the pendulum's motion to simultaneously extract the equilibrium angle and the free-oscillation angular frequency ω_0. In this fitting routine, ω_0 is treated as the single non-linear parameter (seeded by a precise initial estimate derived from the pendulum's period). The difference in equilibrium angles between the two SM positions yields $\Delta\phi_t$. To ensure a traceable measurement of the pendulum period derived from ω_0, the autocollimator data must be precisely timestamped. To achieve this, the autocollimator was modified by the manufacturer to output a TTL signal every time the LED turns on and a measurement is performed (every 40 ms). This TTL signal gates a GPS-disciplined signal generator set at 20 MHz. These clock cycles are counted to provide a highly accurate, traceable relative time base between measurements.

As shown in figure 4, superimposed on the equilibrium change of the pendulum is the free harmonic oscillation. Ignoring dissipation, its frequency obeys

$$\omega_0^2 = \frac{\partial^2 E_{\mathrm{pot}}/\partial \phi_t^2}{I}, \tag{7}$$

where I is the moment of inertia of the torsion pendulum. The total potential comprises (1) the torsion strip (elastic plus gravitational components [5]), (2) coupling to local gravitational fields created by laboratory objects unrelated to the experiment, and (3) coupling to the SMs. At torque extrema, the SM contribution has zero second derivative and does not affect ω_0.

Hence the effective spring constant (from components 1–2) is the same at both positions. Combining equations (5)–(7) gives:

$$G_f = \frac{\omega_0^2 R_s}{16 m_s m_t \Gamma_{\max}} I \Delta\phi_t. \tag{8}$$

The sources of the moment of inertia, whose contributions combine additively, are: (1) the four point masses on the MC radius (MCR) R_t, (2) each TM about its own axis, and (3) the torsion disk and attachments:

$$I = 4 m_t {R_t}^2 + 4 m_t \cdot \frac{1}{2}\left({b_t}^2 - {a_t}^2\right) + I_{\mathrm{disk}}, \tag{9}$$

where b_t and a_t are the outer and inner radii of the cylindrical TMs, respectively. Then, we arrive at the expression

$$\boxed{G_f = \frac{\omega_0^2 R_s R_t^2}{4 m_s \Gamma_{\max}} \left(1 + \frac{b_t^2 - a_t^2}{2R_t^2} + \frac{I_{\mathrm{disk}}}{4 m_t R_t^2}\right) \Delta\phi_t.} \tag{10}$$

with typical corrections $(b_t^2 - a_t^2)/(2R_t^2) = 0.026$ and $I_{\mathrm{disk}}/(4 m_t R_t^2) = 0.11$.

Because the free-deflection method uses inertial torques to calibrate static gravitational torques, the gravitational mass of the TM m_t in the torque equation is effectively canceled by the inertial mass in the equation of motion of the free oscillation. Consequently, the derived value of G_f depends only weakly on m_t. This cancellation is not perfect, however, primarily due to the presence of the torsion disk and other components of the pendulum assembly that contribute differently to the moment of inertia and the gravitational coupling.

3.3. The electrostatic-servo method

Unlike free deflection (where all objects are kept electrically grounded), the servo method grounds the pendulum electrode while driving the opposing electrodes at 1 kHz.

Eight copper cylinders form two opposing groups of four (figure 3); each TM lies between a pair of electrodes. Electrodes of radius 3 mm are positioned to maximize the capacitance gradient with respect to ϕ_t [5, 47], ensuring linear response to first order and no electrostatic stiffness. Driving one electrode group produces a clockwise electrostatic torque, while switching to the opposite group reverses the direction. In the *Mark II* particular attention was paid to avoid the presence of lossy dielectrics in the electric fields between the electrodes. While the mounting insulators were fully shielded to minimize dielectric loss, potential frequency-dependent effects may still arise from the Kapton outer sheaths of the coaxial cabling. Under ac excitation, the oscillation of the electric field experienced by dipoles in such dielectrics produces energy losses not accounted for in the standard electrostatic energy model.

Under quasistatic conditions, the electrostatic energy E_{el} of a system enclosed by a grounded conductor is

$$E_{\mathrm{el}} = \tfrac{1}{2}\mathbf{V}^{\mathsf{T}}\mathbf{C}_{\mathrm{cap}}\mathbf{V}, \tag{11}$$

where $\mathbf{V}$ is the vector of electrode potentials and $\mathbf{C}_{\mathrm{cap}}$ is the symmetric capacitance matrix. For a system enclosed by a grounded conductor, the capacitance matrix of the internal electrodes has vanishing row and column sums. Hence, its diagonal elements satisfy $C_{ii} = -\sum_{i\neq j} C_{ij}$. The cross-capacitance measured by a capacitance bridge is defined as $C_{\mathrm{c},ij} = -C_{ij}$ [47, 48]. Because the system is maintained at constant voltage, the power supply does work to maintain the potentials. Consequently, the electrostatic torque is the positive derivative of E_{el} with respect to ϕ_t, which yields

$$N_{\mathrm{el}} = +\frac{k_{12}}{2}(V_1 - V_2)^2 + \frac{k_{13}}{2}(V_1 - V_3)^2 + \frac{k_{23}}{2}(V_2 - V_3)^2, \text{ where } k_{ij} = \frac{\mathrm{d}C_{\mathrm{c},ij}}{\mathrm{d}\phi_t}. \tag{12}$$

For the torque derivation in equation (12), two critical assumptions are made: (1) the system is completely enclosed by a contiguous grounded conductor (in practice, this is an approximation, as the vacuum vessel requires an optical window for the autocollimator to monitor the pendulum's rotation); (2) the system comprises only three primary electrodes ($n = 3$). However, a rigorous energy calculation must include all conductors within the grounded boundary that possess finite potentials relative to the shields. In particular, the shields of the coaxial cables used to drive the electrodes must be considered. If the cross-capacitances between these cable shields and the primary electrodes are modulated by the motion of the TMs, they contribute an additional, unaccounted-for torque derivative to the total energy sum.

For ac excitation the instantaneous voltages V_1 and V_2 are sinusoidal at angular frequency ω_{ac}, and the torque must be averaged over one cycle.

$$V_1 = V_{1,\mathrm{rms}}\sqrt{2}\sin(\omega_{\mathrm{ac}}t), \tag{13}$$

$$V_2 = V_{2,\mathrm{rms}}\sqrt{2}\sin(\omega_{\mathrm{ac}}t + \alpha), \tag{14}$$

where $V_{1,\mathrm{rms}}$ and $V_{2,\mathrm{rms}}$ are the RMS voltages, ω_{ac} is the drive frequency, and α is the phase difference of the voltages. The pendulum (electrode 3) is at an unknown dc potential V_3 set by contact potentials in the suspension. These spatial variations in surface potential, known as the 'patch effect,' can arise from differences in chemical composition or crystallographic alignment. These potentials are sensitive to vacuum pressure and temperature. However, as long as V_3 remains stable between the two applied torques, its effect vanishes.

Averaging $(V_i - V_j)^2$ over one ac cycle yields

$$\langle V_{12}^2\rangle = V_{1,\mathrm{rms}}^2 + V_{2,\mathrm{rms}}^2 - 2V_{1,\mathrm{rms}}V_{2,\mathrm{rms}}\cos\alpha, \tag{15}$$

$$\langle V_{13}^2\rangle = V_{1,\mathrm{rms}}^2 + V_3^2, \text{ and } \langle V_{23}^2\rangle = V_{2,\mathrm{rms}}^2 + V_3^2. \tag{16}$$

The corresponding average torque is

$$\langle N_{\mathrm{el}}\rangle = \tfrac{1}{2}k_{12}\langle V_{12}^2\rangle + \tfrac{1}{2}k_{13}\langle V_{13}^2\rangle + \tfrac{1}{2}k_{23}\langle V_{23}^2\rangle. \tag{17}$$

The true RMS voltages $V_{1,\mathrm{rms}}$ and $V_{2,\mathrm{rms}}$ are measured with Fluke 5790B ac standards [5] The two 5790B AC measurement standards were calibrated by the Applied Electrical Measurement group at NIST as part of a special test. Their AC voltage traceability is ultimately derived from the NIST multijunction thermal voltage converters. The unknown V_3 cancels when $k_{23} = -k_{13}$, a condition closely met despite small geometric asymmetries.

To linearize the servo response, the RMS voltages are chosen as

$$V_{1,\mathrm{rms}} = V_{\circ,\mathrm{rms}} + V_{\Delta,\mathrm{rms}}, \tag{18}$$

$$V_{2,\mathrm{rms}} = V_{\circ,\mathrm{rms}} - V_{\Delta,\mathrm{rms}}, \tag{19}$$

where $V_{\circ,\mathrm{rms}}$ is the common-mode voltage and $V_{\Delta,\mathrm{rms}}$ the differential component that modulates the torque. With this symmetry, the voltage measured at $\phi_{\max}$ equals that of the opposite electrode at $\phi_{\min}$ (figure 5).

---

[5] Certain equipment, instruments, software, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement of any product or service by the NIST, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

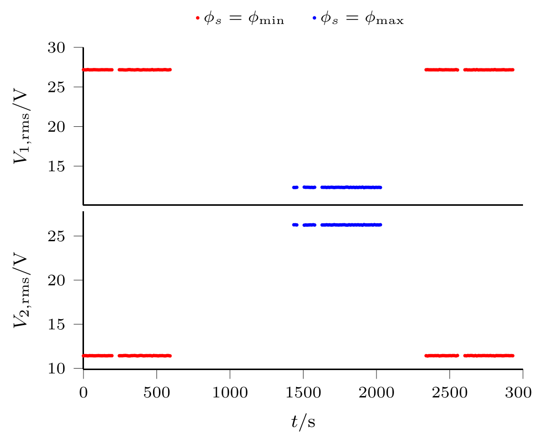

Figure 5. Representative measurements of $V_{1,\mathrm{rms}}$ (top) and $V_{2,\mathrm{rms}}$ (bottom). Red and blue points correspond to SMs at $\phi_{\min}$ and $\phi_{\max}$, respectively. Gaps indicate intervals when the SMs were moved and the servo loop was settling. Missing points arise from convergence issues in the true-RMS voltmeters. Data were acquired with two Fluke 5790B instruments.

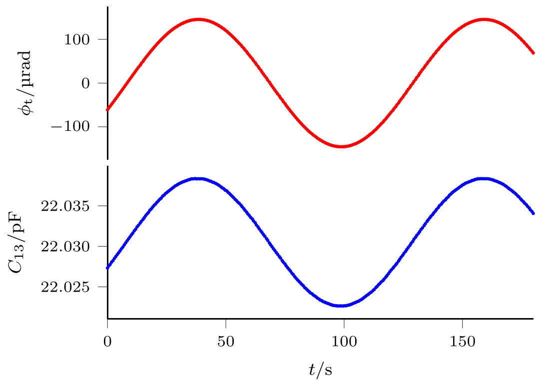

Figure 6. Raw data used to determine the capacitance gradient k_{13}. The top panel shows the torsion angle and the bottom panel the simultaneously recorded cross-capacitance between electrode 1 and the pendulum. Data are collected over 1.5 pendulum periods per run; sine fits yield the gradient. The sequence is repeated five times for each of the three electrode pairs (k_{13}, k_{23}, k_{12}) to assess repeatability.

Each capacitance gradient is measured by exciting the free oscillation of the pendulum to an amplitude of $\phi_a \approx 120\,\mu\mathrm{rad}$ (the voltages are disconnected after initial excitation), while an Andeen Hagerling capacitance bridge (AH2700A) records the cross-capacitance and a Möller-Wedel autocollimator (Elcomat HR) simultaneously records the angular motion. Sine fits to the data yield the gradient. The sequence is repeated five times per electrode pair to assess measurement repeatability (figure 6).

Figure 7 shows the capacitance-gradient data used in the servo-mode determination of G. Four measurement campaigns were conducted: one with sapphire SMs (Sa) followed by three with the SMs (Cu0–Cu2), each in a different rotational orientation. Within each campaign the measured gradients are highly repeatable, while the small offsets between campaigns are consistent with a slight re-seating of the pendulum after the dimensional measurements on the CMM.

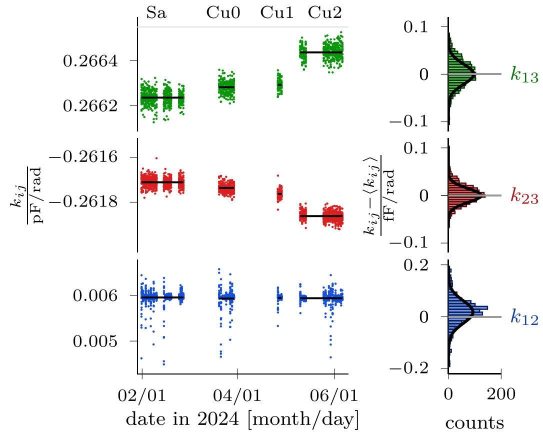

Figure 7. Time evolution of the three mutual capacitance coefficients k_{13}, k_{23}, and k_{12} over the four measurement campaigns. The right column shows the corresponding residual distributions obtained by subtracting each channel's campaign mean.

The relative standard deviations of k_{13} and k_{23} are of the order of 1×10^{-4}, and Allan deviation analysis confirms that these channels can be averaged below 1×10^{-5} given sufficient integration time. The weaker coupling k_{12} exhibits occasional excess noise, likely from a relay in the readout path, but its mean value is stable across all campaigns and its contribution to G is negligible. Because $\langle V_{12}^2\rangle$ is nearly identical in both torque states, its first-order effect cancels, in agreement with equations (18) and (19).

In the servo method, the gravitational constant follows from equation (5) by substituting the electrostatic torque difference,

$$\boxed{G_\mathrm{s} = \frac{R_\mathrm{s}}{16 m_\mathrm{s} m_\mathrm{t} \Gamma_{\max}} \Delta\langle N_\mathrm{el}\rangle,} \tag{20}$$

Equation equation (20) assumes an ideal servo loop with infinite gain. In practice, the controller operates with a finite gain, leaving a small residual twist angle $\Delta\phi_\mathrm{t}$ of the pendulum. To ensure a rigorous torque balance, this residual angle (less than a tenth of an arcsecond) is continuously measured by the autocollimator, and the corresponding elastic torque stored in the torsion strip, $\kappa\Delta\phi_\mathrm{t}$, is added as a small correction to the applied electrostatic torque. The servo algorithm is described in detail in [49].

In the following, we compare the two operating methods and their respective sensitivities.

Table 1. Relative sensitivities of G_f and G_s to the mass and geometry of the experiment.

x	$c_x(G_\mathrm{f})$	$c_x(G_\mathrm{s})$
m_s	-1	-1
m_t	-0.1	-1
R_s	-5.4	-5.4
R_t	2.4	4.4
I_disk	0.11	0
ϕ_t	1	-1

3.4. Comparison of both methods

Beyond the obvious electrical–versus–mechanical distinction, we focus on how each mode depends on mass and geometry. We quantify this with the *relative sensitivity*

$$c_x(y) \equiv \frac{x}{y}\frac{\partial y}{\partial x} = \frac{\partial \ln y}{\partial \ln x}, \tag{21}$$

applied here to $y \in \{G_\mathrm{f}, G_\mathrm{s}\}$ and selected inputs x. Using the results of section 6 for $\partial\Gamma_\mathrm{max}/\partial R_\mathrm{t}$ and $\partial\Gamma_\mathrm{max}/\partial R_\mathrm{s}$, table 1 summarizes $c_x(G_\mathrm{f})$ and $c_x(G_\mathrm{s})$.

The most pronounced differences occur for m_t, R_t, and ϕ_t. The ϕ_t contrast arises because G_f scales with the fitted angle (ϕ_t), whereas the servo's capacitance gradients scale as $1/\phi_\mathrm{t}$.

3.5. Technical data

Table 2 lists the principal experimental parameters.

4. Modifications to the BIPM design

Having described the principle of the experiment, we now discuss the changes made to the original BIPM *Mark II* apparatus and their implications for the measurement.

4.1. Changes to length metrology

The geometry of the *Mark I* apparatus was measured with a DEA Swift CMM and *Mark II* with a Brown & Sharpe Mistral CMM. In both cases, the principal investigator performed the MC measurements. At NIST, the dimensional-metrology group was involved from the outset and operated a Zeiss Xenos CMM.

The switch to the Xenos had three consequences:

- **Thermal environment.** The Xenos operated in an underground lab at 20.00 °C with 0.01 K stability maintained by laminar flow through the gantry. This stability nevertheless introduced a thermal torque (section 7).
- **Active leveling.** During CMM operation the bed was supported by active leveling feet; when parked for G measurements it rested on rubber stops, producing a small change in level and a slow drift between measurements.
- **Protocol.** A new procedure to determine the MCR of the TMs and SMs was developed at NIST (section 5).

4.2. New TMs

To reduce uncertainty in the MCR, the original TMs were replaced by diamond-turned cylinders with a central bore of radius 2.5 mm to mitigate center high-spot form error (section 5). The new masses improve the geometric precision of the R_t determination.

4.3. New torsion disk

The original torsion disk was machined with slightly non-parallel top and bottom surfaces. A new disk was fabricated, and its form was verified by ancillary measurements with the CMM.

4.4. Additional sapphire SMs

Dr. Kibble proposed that the elevated G observed with the BIPM torsion balance could reflect ac magnetic coupling between SMs and TMs [50]; the thin 4 mm chamber wall permits attenuated line-frequency fields to penetrate. Speake's analysis limited the resulting spurious torque to less than 2×10^{-5} of the gravitational signal [51].

To test this conclusively, NIST procured a second set of SMs from single-crystal sapphire with approximately the same dimensions as the copper set. Owing to sapphire's lower density (about 45% of copper's), the gravitational signal was proportionally smaller. The single-crystal uniformity eliminated the azimuthal averaging required for the copper masses.

4.5. New software

We rewrote the entire software suite for instrument control, environmental monitoring, and data analysis at NIST. Special attention was given to redeveloping the mass-integration and moment-of-inertia codes. Because G scales inversely with the mass-integration factor ($G \propto 1/\Gamma_\mathrm{max}$), any relative error in Γ_max transfers directly to the final result. The details of this calculation and its validation are presented in section 5

4.6. Vacuum environment

The same vacuum chamber was used at NIST, but sealing performance was degraded and the pumping geometry (distance to the turbomolecular pump) differed from BIPM's. After observing a pressure dependence of the torque difference, the pumping distance was adjusted (section 7). Following initial pump-down, the pressure rose slowly with a time constant of about one month (e.g. 7.4×10^{-3} Pa to 1.7×10^{-2} Pa over two months), behavior that ultimately revealed the thermal gas effect (section 7).

4.7. Balancing balls

Because the suspended system's center of mass was offset from the suspension point, brass balancing spheres with a total mass of 29.196 g were placed in

Table 2. Technical parameters of the experiment. For each of the three mass sets, values are averaged over the four masses.

Category	Description	Expression	Value	
TMs	Mass	m_t	1150.156	g
	Inner radius	a_t	2.502	mm
	Outer radius	b_t	27.481	mm
	Height	h_t	54.910	mm
	MC radius	R_t	120.0315	mm
Copper SMs	Apparent mass[a]	m_s	11 191.68	g
	Outer radius	b_s	58.987	mm
	Height	h_s	115.008	mm
	MC radius[b]	R_s	213.9642	mm
	Peak-to-peak torque[b]	ΔN_max	31.1979	nN m
Sapphire SMs	Apparent mass[a]	m_s	5015.60	g
	Outer radius	b_s	59.015	mm
	Height	h_s	115.359	mm
	MC radius	R_s	213.9642	mm
	peak-to-peak torque	ΔN_max	13.9799	nN m
Torsion balance	Moment of inertia (disk)	I_disk	7.45×10^{-3}	kg m^2
	Pendulum period (with TMs)	T_0	121	s
Cross capacitances between electrodes	One and two	C_{12}	5.6	pF
	One and pendulum	C_{13}	22.03	pF
	Two and pendulum	C_{23}	23.4	pF
Capacitance gradients	One and two	k_{12}	−0.006	pF rad^{-1}
	One and pendulum	k_{13}	0.262	pF rad^{-1}
	Two and pendulum	k_{23}	−0.266	pF rad^{-1}

[a] Apparent masses are corrected for air buoyancy.

[b] Correspond to the 0-mark-out rotation ($\theta = 0$).

selected holes on the torsion disk to level it; several sizes and placements were tested. Their contribution was included in both Γ_max and I_disk. For *Mark II* at BIPM three balancing balls were used with total mass of 22.15 g. In the NIST configuration the torsion strip was oriented so the autocollimator viewed it edge-on.

5. Dimensional metrology

Of the changes described above, dimensional metrology is the most significant and is treated in detail below. Measurements of R_s and R_t were performed using a Zeiss Xenos CMM with a working volume of 900 mm × 1500 mm × 700 mm. A tactile probe with a 296 mm long, ruby-tipped stylus was used to acquire points on the surface of the TMs and SMs. The tops and sides of the masses were measured. Least-squares fits were applied to construct the reference planes and cylinders, respectively. The geometric center of the cylinder is calculated by projecting the planes down the axis of the cylinder to a distance that is half the cylinder height and calculating that intersection point. The four intersection points obtained from each set of masses are used to calculate a MCR for R_s and R_t with a least-squares method. The cylinder heights were measured with each cylinder lying on its side to determine the mean distance between the top and bottom faces.

Unfortunately, the tactile probe did not have complete azimuthal access around the masses due to the mechanical constraints of the torsion balance assembly. Access to the SMs was restricted by the inner ledge of the vacuum chamber base, under which the masses are partially recessed. Access to the TMs was even more limited due to the presence of the overhead yoke (the silver structural component visible in figure 8). These physical barriers prevented the CMM probe from traveling a full 360° path, limiting measurement to an arc of approximately 120° for the TMs and 190° to 210° for the SMs. The masses used at the BIPM and originally supplied to NIST exhibit a tri-lobed form error at the micrometer level, and the SMs are slightly tapered. With full 360° access, a best-fit cylinder would average the lobes and taper symmetrically, reducing bias in the inferred center and diameter. When only a partial arc is measured, however, fitting a cylinder to a tri-lobed shape biases the estimated center and diameter (see figure 9), which in turn biases R_s and R_t [52]. The combination of partial coverage and 10 μm to 20 μm form error therefore increases the measurement uncertainty beyond what is desired. Mitigation strategies are presented below.

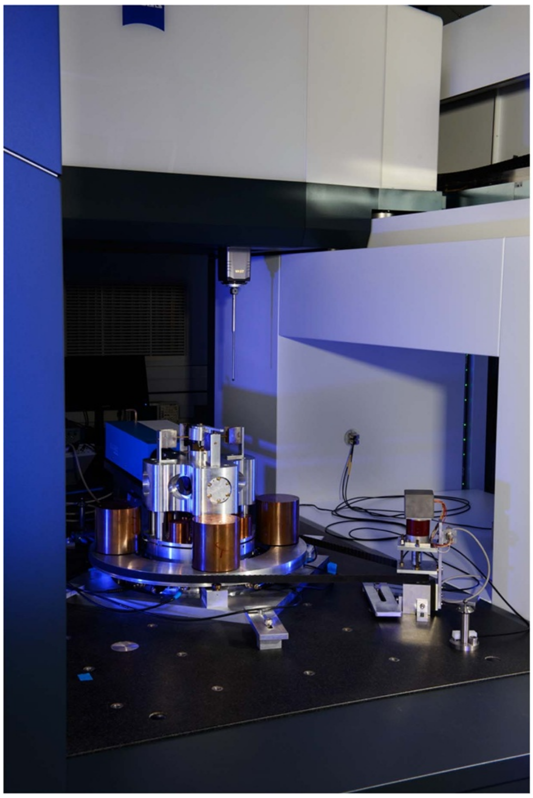

Figure 8. The BIPM torsion balance mounted on the Zeiss Xenos coordinate-measuring machine. The vacuum chamber was removed for this photo. Photo by Jennifer Lee (NIST).

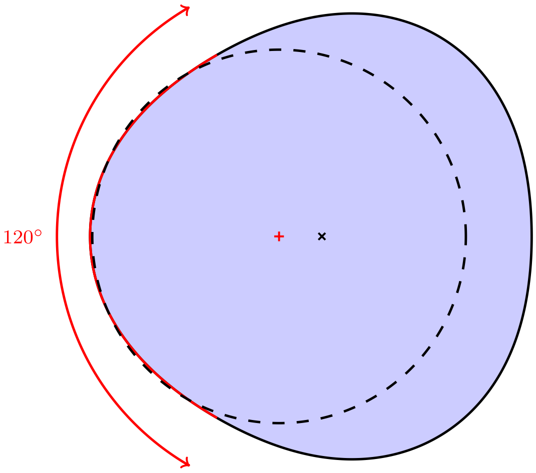

Figure 9. Difference between best-fit centers obtained from partial-arc (+) and full-coverage (×) CMM data for a cylinder exhibiting a tri-lobed form error.

New TMs were manufactured by single-point diamond turning achieving cylindricity form error below 0.25 µm. By improving the cylindrical form of these masses, the difference between the two estimates of the MCR resolved from partial-coverage and full-coverage data is about 150 nm.

Table 3. Sensitivity of R_s and R_t to CMM angular coverage. Comparison is made between partial coverage (190° to 210° for SMs; 120° for TMs) and full 360° coverage. For the copper source masses, the deviation depends on the carousel position ϕ_s and the orientation θ. The difference is defined as $\Delta R_\mathrm{sm} \equiv R_\mathrm{partial} - R_\mathrm{full}$ and listed in µm.

Mass (type, pos orientation)	$\Delta R_\mathrm{sm}/\mu\mathrm{m}$
SM, copper$\theta = 0°$, $\phi_\mathrm{s} = \phi_\mathrm{max}$	-0.5
SM, copper$\theta = 120°$, $\phi_\mathrm{s} = \phi_\mathrm{max}$	0.6
SM, copper$\theta = 240°$, $\phi_\mathrm{s} = \phi_\mathrm{max}$	-1.1
SM, sapphire $\phi_\mathrm{s} = \phi_\mathrm{max}$	1.4
SM, copper$\theta = 0°$, $\phi_\mathrm{s} = \phi_\mathrm{min}$	-0.3
SM, copper$\theta = 120°$, $\phi_\mathrm{s} = \phi_\mathrm{min}$	-0.2
SM, copper$\theta = 240°$, $\phi_\mathrm{s} = \phi_\mathrm{min}$	-0.3
SM, sapphire $\phi_\mathrm{s} = \phi_\mathrm{max}$	2.3
TM, original	0.98
TM, diamond turned	-0.05

5.1. Determination of the test-mass MC

The parameter R_t is measured with the tactile probe of the CMM to measure 72 points on a partial cylindrical surface. This surface covers about 120° circumferentially and 40 mm axially with the points equally distributed. Using the measured points, a constrained least-squares fit is used to calculate the cylinder center locations and R_t. Applying a constrained-diameter least-squares fit requires a fixed diameter for each of the TMs. To determine the constrained diameter to use in the constrained fit (specific for each TM), the TMs were measured at times when the TMs were removed from the torsion disk. Each TM was measured with a full 360° of coverage using the Xenos CMM, and a least-squares fit was employed to determine each diameter.

Over the past five years of measuring, the value of R_t was about 120.0315 mm, and it has drifted by about 1.5 µm over that time. The long term drift is not an uncertainty source, since each G measurement campaign obtains the R_t value by measurement for its own (much smaller) time period.

The old TMs showed a difference in R_t between the measurement with full and partial coverage of 0.98 µm, see table 3. After we exchanged the original TMs with diamond-turned ones, the value of R_t is nearly unaffected by the limited sampling. Once all measurement campaigns were completed, R_t was measured using both partial and 360° of coverage in a single setup to verify that the difference in these values was only about ≈ 50 nm using the two different measurement strategies.

5.2. Determination of the source-mass MC

The SMs rest on the carousel near the bottom of the device, and have a nominal MCR of 213.9642 mm. Unlike the TMs, the SMs have better CMM access, although the access varies with carousel angle, ϕ_s, because of nearby vacuum plumbing and chamber port covers. The available CMM coverage around a

Table 4. Standard uncertainties ($k = 1$) contributing to the determination of R_s and R_t. The uncertainties are rounded to the nearest 10 nm.

Source	$u(R_s)$/nm	$u(R_t)$/nm
CMM error[a]	150	150
Form error[b]	100	100
Temperature	140	80
CTE	40	20
Combined	**230**	**200**

[a] includes repeatability.

[b] includes form error.

SM typically ranges from 190° to 210° depending on the particular SM and ϕ_s.

Similar to the case of the TMs, we sample a partial cylindrical surface patch of each SM with the CMM's tactile probe, using 44 points uniformly distributed on the patch. And just like the TMs, measuring just a partial arc of a cylinder, especially one with imperfect form, would result in an error in determining its location, and thus an error in the MCR obtained.

After all the G measurements were concluded, the entire torsion balance device was removed except for the carousel and the SMs. In this state, a series of experiments were performed using a measurement plan that probes points all the way around each SM, which allows comparison of the results to the original measurement plan that only partially probes each cylinder. With different measurement schemes, a difference in the MCR is likely to be present. These experiments help determine that difference and would allow corrections to be made to the CMM measurement results.

Measurements all the way around each SM utilize the same point spacing as the scheme used during the G experiments, but with an increase to the number of probing points to 72. The typical differences between these two measurement strategies for the resultant MCR are summarized in table 3.

5.3. Measurement uncertainty for the MC radii

Table 4 summarizes the dominant sources of measurement uncertainty for R_s and R_t. These are discussed below.

The CMM determines the MC radius required for the $\Gamma_{\rm max}$ calculation and applies a correction for the CTE, $\alpha_{\rm T}$. This correction is proportional to $R\alpha_{\rm T}\Delta T$, where R is the measured radius and ΔT is the temperature difference from 20.00 °C.

5.3.1. Uncorrected CMM error

To quantify the CMM contribution to the measurement uncertainty in R_s and R_t, we include (i) the CMM's length-measurement performance and (ii) the repeatability of the R_s and R_t determinations. Although the CMM underwent volumetric error mapping at the start of the measurement campaign, its accuracy can drift over time, particularly in response to large, rapid temperature excursions in the laboratory during scheduled outages in the spring and fall. To monitor accuracy, we used a step gauge made by KOBA that had been independently and traceably calibrated on a higher-accuracy CMM [53]. After the campaign, the Xenos CMM was evaluated using this gauge; the results are summarized in figure 10.

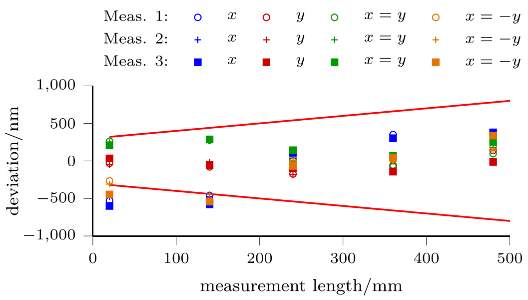

Figure 10. Step-gauge verification of the Xenos CMM. A step gauge was measured in four in-plane orientations, X, Y, $X = Y$ (+45°), and $X = -Y$ (−45°) with three repeated runs in each orientation. The red lines indicate the maximum permissible error (MPE) envelope. For measurement lengths $\geqslant 240$ mm, all deviations lie within the MPE.

We place the step gauge in the same measurement volume as the torsion balance. The step gauge is oriented along the X, Y, and the two diagonals (±45°) in the XY plane. The red lines in figure 10 show the CMM's MPE as a function of length. In our evaluation, deviations exceeding the MPE were observed for lengths below 240 mm; at and above 240 mm all readings were within the specification. The features used to determine R_s and R_t involve lengths $\geqslant 240$ mm, so we adopt an MPE-consistent bound of $\pm 0.5\,\mu$m on the measured MCR reported by the CMM. Modeling this CMM error with a uniform probability distribution function, the estimated standard uncertainty for measuring the diameter of the SM and TM MC is about $0.5\,\mu\text{m}/\sqrt{3} = 0.29\,\mu$m.

When measuring the SMs and TMs, the standard deviation of resolving the MCR is about 0.1 µm for both mass sets. Summing this value and the standard deviation from CMM measurement error in quadrature yields a standard uncertainty of about 0.3 µm for the diameter or 0.15 µm for the radius.

5.3.2. Imperfect cylinder form

Unlike the diamond-turned TMs, the form of the SMs is not so perfect. Although the effect of azimuthal coverage on the measurement has been investigated and corrected, see table 3, uncertainties remain. The best estimate for this uncertainty is approximately the standard deviation of resolving the MCR using the 360 degree measurement data from table 3, which is about 0.1 µm.

5.3.3. Effect of imperfectly known temperature

The CMM has the ability to correct for thermal expansion due to temperature fluctuations of the parts and reports the SM and TM MCR as if it was measured at 20.00 °C. However, there are small temperature fluctuations and non-uniformity while the measurements are occurring. As a result there will be an error in making this correction. Both the SMs and TMs are supported on aluminum platters that have nominal radii of 214 mm and 120 mm, respectively. During CMM operation, the measured temperature of the torsion balance is about 19.90 °C. With aluminum having a nominal CTE of $\alpha_\mathrm{T} = 22.5 \times 10^{-6}\,\mathrm{K}^{-1}$, the maximum expected change in MC radius for the SM and TM cluster is 0.48 µm and 0.27 µm, respectively. These values represent the magnitude of the deterministic correction. The residual uncertainty of this correction, arising from the temperature non-uniformity and the uncertainty of the CTE, is treated as a separate uncertainty source and is discussed below.

During measurements the temperature fluctuates and has a non-uniformity of about ±0.05 K. Applying a uniform probability distribution function to this temperature variability, a standard uncertainty of 0.03 K is estimated (calculated as $0.05\,\mathrm{K}/\sqrt{3}$).

5.3.4. CTE

There is uncertainty in the CTE which can induce uncertainty in measuring R_s and R_t. The aluminum used to construct the torsion disk is MIC-6 aluminum while the carousel is an unknown alloy. Based on published values, α_T for various grades of common commercial off-the-shelf aluminum (MIC-6, 6061, 2024, and 7075) ranges between $21 \times 10^{-6}\,\mathrm{K}^{-1}$ to $24 \times 10^{-6}\,\mathrm{K}^{-1}$. Applying a uniform probability distribution to this range, a standard uncertainty of $1.7 \times 10^{-6}\,\mathrm{K}^{-1}$ is obtained (calculated as $3 \times 10^{-6}\,\mathrm{K}^{-1}/\sqrt{3}$).

6. Mass integration

With the dimensional metrology established, we now describe how this information is used in the analysis. The mass-integration constant, Γ, is a dimensionless geometry factor that relates the measured torque N on the torsion balance to the gravitational constant G via

$$N(\Delta\phi) = G\frac{8m_\mathrm{s}m_\mathrm{t}}{R_\mathrm{s}}\Gamma(\Delta\phi), \tag{22}$$

This section describes the computation of the mass-integration constant Γ. All calculations in this section use the masses and dimensions listed in table 5. Importantly, the mass value m_s represents the *apparent mass* (weighed in air) rather than the true vacuum mass; this implicitly and exactly subtracts the gravitational contribution of the air displaced when the SMs are moved between measurement positions. Before discussing the numerical mass integration, we present an analytical model useful for uncertainty propagation.

Table 5. Parameters used for the calculation of Γ_max in this section. Values are averages across the four masses in each set.

par.	Description	Value	
m_t	Mass of TM	1150.156	g
h_t	Height of TM	54.910	mm
a_t	Bore radius in TM	2.502	mm
b_t	Outer radius of TM	27.481	mm
m_s	Mass of SM	11 191.68	g
h_s	Height of SM	115.008	mm
b_s	Radius of SM	58.987	mm
R_t	MCR TMs	120.0315	mm
R_s	MCR SMs	213.9642	mm

6.1. Simplified analytical equation for the torque

Let all masses be idealized as point masses. The gravitational potential between a single TM at cylindrical coordinates $(R_\mathrm{t}, \phi_\mathrm{t})$ and a SM at $(R_\mathrm{s}, \phi_\mathrm{s})$, both at $z = 0$, is given by

$$V(\Delta\phi) = -G\frac{m_\mathrm{s}m_\mathrm{t}}{\sqrt{{R_\mathrm{t}}^2 + {R_\mathrm{s}}^2 - 2R_\mathrm{t}R_\mathrm{s}\cos(\Delta\phi)}}. \tag{23}$$

The potential depends on the angle difference $\Delta\phi = \phi_\mathrm{s} - \phi_\mathrm{t}$.

The inverse distance can be expanded in a Legendre series:

$$\frac{1}{\sqrt{{R_\mathrm{t}}^2 + {R_\mathrm{s}}^2 - 2R_\mathrm{t}R_\mathrm{s}\cos\Delta\phi}} = \frac{1}{R_\mathrm{s}}\sum_{l=0}^{\infty}\left(\frac{R_\mathrm{t}}{R_\mathrm{s}}\right)^{l} P_l(\cos\Delta\phi). \tag{24}$$

Including the remaining three SMs at $(R_\mathrm{s}, \phi_\mathrm{s} + n\pi/2)$ for $n \in \{1, 2, 3\}$ yields

$$V(\Delta\phi) = -G\frac{2m_\mathrm{s}m_\mathrm{t}}{R_\mathrm{s}} \times \sum_{l=0}^{\infty}\left(\frac{R_\mathrm{t}}{R_\mathrm{s}}\right)^{2l}\left(P_{2l}(\cos\Delta\phi) + P_{2l}(\sin\Delta\phi)\right). \tag{25}$$

Only Legendre polynomials of even degree $(2l)$ contribute due to the fourfold azimuthal symmetry of the SMs. Accounting for the four TMs, located at $(R_\mathrm{t}, \phi_\mathrm{t} + n\pi/2)$ amounts to multiplying equation (25) by four. We note that the sum depends on the ratio of $R_\mathrm{t}/R_\mathrm{s}$, which we will abbreviate as η.

The resulting torque is given by the derivative of the full potential with respect to $\Delta\phi$:

$$N(\Delta\phi) = G\frac{8m_\mathrm{s}m_\mathrm{t}}{R_\mathrm{s}}\Gamma(\Delta\phi,\eta) \tag{26}$$

and the function

$$\Gamma(\Delta\phi,\eta) = \sum_{l=1}^{\infty}\eta^{4l}\left(P_{4l}^{1}(\cos\Delta\phi) - P_{4l}^{1}(\sin\Delta\phi)\right). \tag{27}$$

The terms P_{4l}^{1} denote associated Legendre polynomials[6] of order one and degree $4l$. Differentiating equation (25) removes all contributions with l odd, because the two terms in brackets cancel in that case. Thus, from the even sequence $2l$ that survives in the potential, only every second term contributes to the torque, leaving degrees divisible by four.

The measurement is carried out at an angle difference $\Delta\phi = \Delta\phi_\mathrm{max}$ where the torque has an extremum. Experimentally the extremum is found by dithering the SM position. At $\Delta\phi_\mathrm{max}$ the torque is

$$\Gamma_\mathrm{max} = \Gamma(\Delta\phi_\mathrm{max},\eta). \tag{28}$$

Using the sensitivity introduced in equation (21), $c_\eta(\Gamma_\mathrm{max})$ can be calculated as follows

$$\frac{\eta}{\Gamma_\mathrm{max}}\frac{\partial\Gamma_\mathrm{max}}{\partial\eta} = \eta\frac{\sum_{l=1}^{\infty}(4l)\,\eta^{4l-1}\alpha_l}{\sum_{l=1}^{\infty}\eta^{4l}\alpha_l} \approx 4.58. \tag{29}$$

Truncating the series at $l = 1$ yields $c_\eta(\Gamma_\mathrm{max}) = 4$, but including higher order terms causes the series to converge to 4.58. This strong sensitivity implies that even a modest relative uncertainty in η, for example 2.1×10^{-6}, propagates into approximately 10×10^{-6} relative uncertainty in Γ_max and consequently in G. The value of Γ_max for the eight point masses is 0.457 946.

6.2. Numerical calculation

The Γ_max calculated from point masses is relatively smaller by 5.9×10^{-2}, see table 10, than a numerical integration over all eight cylinders. Thus, Γ_max must also depend on the dimensions of the cylinders. Assuming all SMs and TMs are identical, Γ_max depends on eight parameters,

$$\Gamma_\mathrm{max}(\eta, a_\mathrm{t}, b_\mathrm{t}, h_\mathrm{t}, b_\mathrm{s}, h_\mathrm{s}, \Delta z, \Delta r), \tag{30}$$

where a_t and b_t are the inner and outer radii of the TMs, b_s is the outer radius of the SM, and h_t and h_s are their heights, respectively. The vertical and horizontal differences between the centers of the TM and SM assemblies are denoted by Δz and Δr. For the nominal parameters, given in table 2, $\Gamma_\mathrm{max} = 0.485099$.

Table 6. Uncertainty budget for the calculation of Γ_max. The column $c_x(\Gamma_\mathrm{max})$ represents the relative sensitivities of Γ_max to the input parameters, derived via numerical integration. Rows are sorted by the absolute magnitude of the sensitivities. The 'Trust' entry reflects confidence in the mass integration (see section 6.8).

x	$u(x)/x$	$c_x(\Gamma_\mathrm{max})$	$(\Gamma_\mathrm{max})/\Gamma_\mathrm{max} \times 10^6$
η	2.6×10^{-6}	4.4	11.3
h_s	8.7×10^{-6}	-3.6×10^{-1}	3.1
b_s	7.6×10^{-6}	1.7×10^{-1}	1.3
h_t	1.8×10^{-5}	-5.4×10^{-2}	1.0
b_t	1.1×10^{-5}	4.0×10^{-2}	0.4
R_s	2.1×10^{-6}	-1.2×10^{-2}	0.0
a_t	8.0×10^{-4}	3.3×10^{-4}	0.3
Trust	1.0	1.1×10^{-5}	11.0
	combined		16.2

The numerical integration algorithms used here [54] were cross-checked against those used at BIPM. Both calculated identical torques for identical input features.

Due to symmetry, Γ_max depends quadratically on Δz and Δr. We find

$$\frac{\Gamma_\mathrm{max}(\Delta z,\Delta r)}{\Gamma_\mathrm{max}(\Delta z = 0,\Delta r = 0)} - 1 = \left(\frac{\Delta r}{1\,\mathrm{mm}}\right)^2 5.34 \times 10^{-5} - \left(\frac{\Delta z}{1\,\mathrm{mm}}\right)^2 1.07 \times 10^{-4}. \tag{31}$$

Thus, the dependence of the mass integration factor on concentricity and vertical alignment is rather weak. One tenth of a millimeter changes Γ_max relatively by 0.5×10^{-6} and -1.1×10^{-6}.

The relative sensitivities of the other six parameters were computed numerically. They are listed in table 6. Two points are worth noting: (1) Numerically, $c_\eta(\Gamma_\mathrm{max})$ is 4.44, only 3% below the analytic point-mass value of 4.58. (2) Theoretically, $c_{R_\mathrm{s}}(\Gamma_\mathrm{max}) = 0$. Our numerical model shows a very weak dependence of Γ_max on the source-mass MCR R_s even with η held fixed. This very weak dependence is negligible, because the factor $8m_\mathrm{s}m_\mathrm{t}/R_\mathrm{s}$ has a relative sensitivity coefficient of -1 dominating in the final result.

6.3. Two useful differential identities

The mass-integration factor has no explicit partial derivative with respect to R_t, because the dependence of Γ_max on R_t is contained implicitly in its dependence on η. It becomes apparent in the total derivative:

$$\frac{\mathrm{d}\Gamma_\mathrm{max}}{\mathrm{d}R_\mathrm{t}} = \frac{\partial\Gamma_\mathrm{max}}{\partial\eta}\frac{\partial\eta}{\partial R_\mathrm{t}} = 4.4\,\frac{\Gamma_\mathrm{max}}{R_\mathrm{t}}. \tag{32}$$

Thus,

$$\frac{R_\mathrm{t}}{\Gamma_\mathrm{max}}\frac{\mathrm{d}\Gamma_\mathrm{max}}{\mathrm{d}R_\mathrm{t}} = 4.4. \tag{33}$$

---

[6] Some Python libraries only evaluate $P_\ell^1(x)$ correctly for $x \in [0,1]$. Symmetry properties of the associated Legendre functions can be used to extend evaluation to other domains, $P_{4l}^1(x) = (-1)^{4l+1}P_{4l}^1(-x)$.

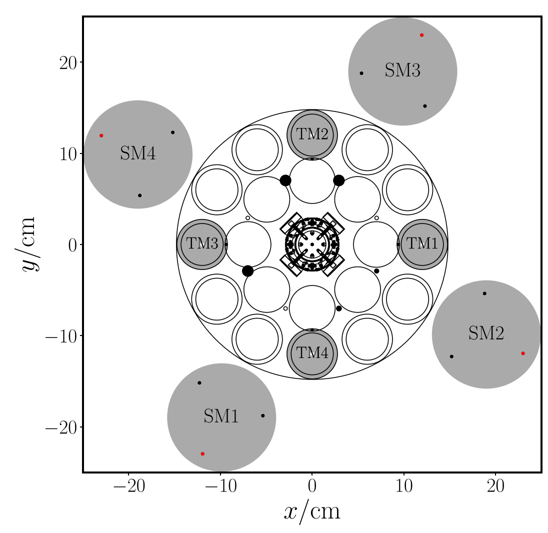

Figure 11. Top view of the complete geometric model used for mass integration. The SMs (SM1–SM4) and TMs (TM1–TM4) are shaded gray. Black and red dots within the SMs mark the kinematic mounting points; the red dots indicate the orientation of each source mass. All visible features, including cutouts and mounting interfaces, are included in the numerical calculation. The solid black circles on the torsion disk denote the balancing balls used for leveling (see section 4.7).

An analogous relation holds for the MCR of the SMs:

$$\frac{R_\mathrm{s}}{\Gamma_\mathrm{max}}\frac{\mathrm{d}\Gamma_\mathrm{max}}{\mathrm{d}R_\mathrm{s}} = -4.41. \tag{34}$$

6.4. Non-identical masses

In practice, the SMs and TMs are not perfectly identical. Accounting for these asymmetries increases Γ_max by a relative amount of -5.2×10^{-7}.

6.5. Other components

So far, we have considered only the SMs and TMs. The torsion balance, however, contains additional components that also contribute to the mass-integration constant. We collectively refer to these components as the 'disk,' i.e. the torsion bob excluding the TMs.

Figure 11 provides an overview of the mostly complete horizontal layout used in the integration. It shows the relative placement of the four SMs and four TMs, as well as the surrounding disk structure and alignment features. This top view also highlights the large lever arms that dominate the torque, motivating the need to model the source-mass geometry in full detail.

Including the disk increases Γ_max by a relative amount of 1.044×10^{-3} (see table 10). The share of TM and disk contribution to Γ_max is shown in table 7. The location and size of all geometric features of the disk were measured at NIST and were found to agree with the drawings furnished by BIPM, except for one blind hole whose contribution to the mass integration is insignificant.

Table 8 further resolves the *signed* contributions of the disk by geometric shape used in the mass integration; entries are sorted by $|\Gamma_\mathrm{max}/\Gamma_\mathrm{max,disk}|$. The dominant term arises from vertical cylinders, which are straightforward to model. To reach a relative accuracy better than 1×10^{-5}, it suffices to include only the first four shape types: vertical cylinders, footballs (the lens-shaped intersection of two equal cylinders whose axes are parallel), horizontal cylinders, and spheres. This hierarchy reflects the strong dependence of gravitational torque on radial position (e.g. for TM terms, the torque grows strongly with the radial position, $\sim R^4$). Most of the more complicated shapes lie near the center of the torsion bob, where their influence is strongly suppressed. The disk contributes only about 0.12% to Γ_max (table 7). A complementary radial view of the integration domain is shown in figure 12.

6.6. Clockwise and counterclockwise torque

In the experiment, we can only measure torque differences with the SMs being advanced clockwise versus counterclockwise from the TMs. We denote the torque in the clockwise configuration as N_max and in the counterclockwise configuration as N_min. The torque difference is defined as

$$\Delta N = N_\mathrm{max} - N_\mathrm{min}. \tag{35}$$

Correspondingly, we calculate the mass integration difference

$$\Delta\Gamma = \Gamma_\mathrm{max} - \Gamma_\mathrm{min}. \tag{36}$$

For comparison to single-state calculations we define the per-state effective constant $\Gamma_\mathrm{eff} = \Delta\Gamma/2$. In the perfectly symmetric case $\Gamma_\mathrm{eff} = \Gamma_\mathrm{max}$. However, small asymmetries in the disk, for example, caused by the balancing balls, produce the observed relative offset of 7.3×10^{-5}.

6.7. Inhomogeneous SMs

The calculations above were conducted with uniform-density (homogeneous) SMs. While this assumption holds for the single-crystal sapphire SMs, small but detectable inhomogeneities were observed in the copper ones [55]. The most likely inhomogeneity is a linear density gradient along the diameter of the SM. This gradient was corroborated by hydrostatic-weighing measurements of samples cut from the same copper ingot used to machine the BIPM SMs. The report by the BIPM team also rules out the presence of millimeter-sized voids based on x-ray analysis of the SMs.

Figure 13 shows the top view of a SM with a density gradient. The density within the SM is modeled as:

$$\rho(x,y,z) = \rho_0 + \Delta\rho\left(\frac{x}{R_\mathrm{s}}\cos\alpha_\mathrm{fid} + \frac{y}{R_\mathrm{s}}\sin\alpha_\mathrm{fid}\right), \tag{37}$$

Table 7. Absolute contributions of the TMs and the disk.

Shape	Count	$\text{mass}\,\text{g}^{-1}$	$I/(\text{g cm}^2)$	I/I_{total}	$\Gamma_{\max}/\Gamma_{\text{total}}$
Test masses	8	4600.625	680 354	0.9014	0.9988
Disk	299	1792.911	74 403	0.0986	0.0012
Total	307	6393.536	754 757	1.0000	1.0000

Table 8. Relative contributions of the individual shapes to the disk component of the mass integration (cf table 7). Entries are sorted by the absolute value of $\Gamma_{\max}/\Gamma_{\max,\text{disk}}$. A football is the lens-shaped intersection of two equal cylinders whose axes are parallel.

Shape	Count	m/m_{disk}	I/I_{disk}	$I/(4m_\text{t}R_\text{t}^2)$	$\Gamma_{\max}/\Gamma_{\text{disk}}$	$\Gamma_{\max}/\Gamma_{\text{total}}$
Vert. cyl.	110	7.39×10^{-1}	9.27×10^{-1}	1.04×10^{-1}	1.04	1.29×10^{-3}
Football	4	8.31×10^{-4}	1.76×10^{-3}	1.98×10^{-4}	8.52×10^{-2}	1.05×10^{-4}
Hor. cyl.	44	1.30×10^{-1}	3.45×10^{-2}	3.88×10^{-3}	-6.56×10^{-2}	-8.10×10^{-5}
Sphere	5	1.63×10^{-2}	2.27×10^{-2}	2.55×10^{-3}	-5.65×10^{-2}	-6.98×10^{-5}
Prism	44	1.21×10^{-1}	1.38×10^{-2}	1.54×10^{-3}	-5.77×10^{-3}	-7.13×10^{-6}
Vert. semicyl.	8	-1.46×10^{-3}	-1.50×10^{-4}	-1.68×10^{-5}	3.14×10^{-4}	3.88×10^{-7}
Pie	8	-1.43×10^{-2}	-5.32×10^{-4}	-5.97×10^{-5}	-8.26×10^{-5}	-1.02×10^{-7}
Point mass	56	1.17×10^{-2}	1.82×10^{-3}	2.04×10^{-4}	5.81×10^{-5}	7.18×10^{-8}
Hor. semicyl.	8	-1.19×10^{-3}	-1.66×10^{-4}	-1.86×10^{-5}	7.28×10^{-8}	8.99×10^{-11}
Cone	12	-1.52×10^{-3}	-2.91×10^{-4}	-3.27×10^{-5}	-1.09×10^{-10}	-1.34×10^{-13}
Disk	299	1.00	1.00	1.12×10^{-1}	1.00	1.23×10^{-3}

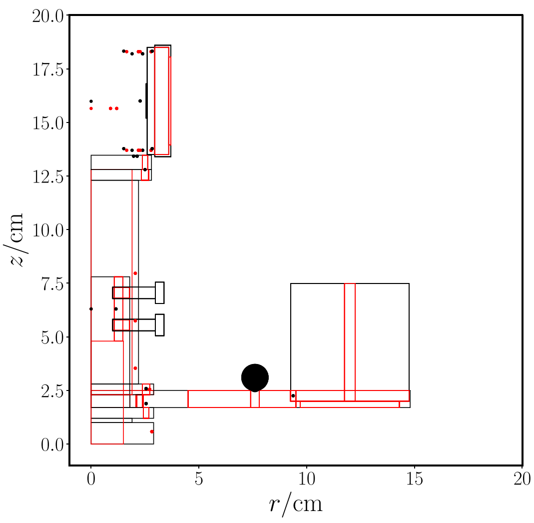

Figure 12. Radial cross-sectional projection of the mass-integration model onto the (r, z) plane. Objects outlined in black correspond to positive-density material, while those in red denote zero-density regions (cutouts or cavities). Small dots mark items represented as point masses in the model.

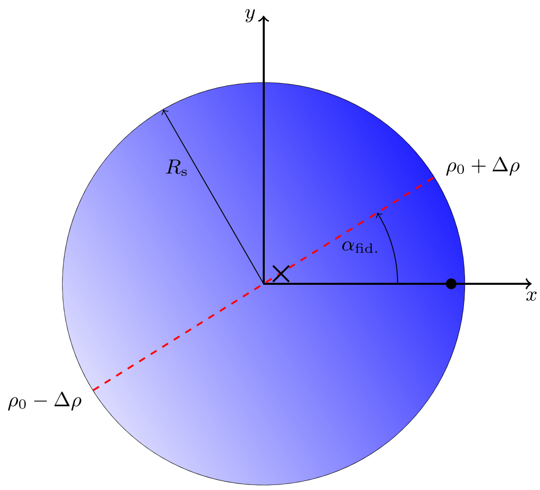

Figure 13. Top view of a source mass relative to a local coordinate system; the x axis points toward a fiducial mark on the mass (black circle). A density gradient runs along the red dashed line; α_{fid} denotes the angle of the direction of increasing density. The geometric center lies at the coordinate origin, but the center of mass is shifted along the gradient, marked by x (exaggerated for clarity).

where α_{fid} is relative to a fiducial mark on the mass. We convert α_{fid} to the laboratory-fixed coordinate system and calculate shifts of the center of mass in the x and y directions:

$$\Delta x_{\text{lab}} = \frac{R_\text{s}}{4}\frac{\Delta\rho}{\rho_0}\cos\alpha_{\text{lab}}, \qquad \Delta y_{\text{lab}} = \frac{R_\text{s}}{4}\frac{\Delta\rho}{\rho_0}\sin\alpha_{\text{lab}}. \tag{38}$$

The center of mass is displaced by one quarter of the fraction $\Delta\rho/\rho_0$ of the source radius along the gradient direction in the laboratory, given by α_{lab}.

The shift of the center of mass relative to the geometric center can be detected by floating the SM in an air bearing and observing the oscillation about its equilibrium position orientation, defined by α_ρ. The period of oscillation is given by:

$$T_\text{s} = 2\pi\sqrt{\frac{2R_\text{s}}{g}\frac{\rho_0}{\Delta\rho}}. \tag{39}$$

The BIPM team performed these experiments for all four SMs and provided values for T_s and α_{lab} which are reproduced in table 9.

Table 9. Provided measurements of the oscillation period and equilibrium orientation of the four SMs in the air bearing. The angles are given in the laboratory-fixed coordinate system. These angles correspond to the mark 1-out rotation ($\theta = 120°$). These measurements allow determination of the relative density gradient and the associated center-of-mass shift.

#	T_s/s	$\alpha_{\text{lab}}/°$	$\frac{\Delta\rho}{\rho_0}$ $\times 10^6$	Δx_{lab} µm	Δy_{lab} µm
1	68.0	111	103	−0.54	+1.41
2	58.8	15	137	1.96	+0.52
3	54.4	291	160	0.85	−2.21
4	44.7	40	238	2.68	+2.25

Given the measured density gradient, we numerically integrate over the SMs to account for the inhomogeneity. Each SM can be rotated in 120° increments by means of the kinematic mount, and all four masses are rotated simultaneously in a clocked configuration. The corresponding orientation angle is denoted by θ. We then compute the resulting torque $N_{\text{inh}}(\theta)$ for the inhomogeneous masses and compare it with the torque produced by homogeneous ones, assuming the masses are at the same R_s.

$$N_{\text{inh}}(\theta) = N_{\text{hom}} + N_a \sin(\theta - \theta_0)\,, \tag{40}$$

with $\theta_0 = 27°$ and $N_a/N_{\text{hom}} = 2.4 \times 10^{-5}$. Because the experiment measures all three clocked orientations, the sinusoidal contribution averages to zero:

$$N_{\text{hom}} = \frac{1}{3}\left(N_{\text{inh}}(0°) + N_{\text{inh}}(120°) + N_{\text{inh}}(240°)\right). \tag{41}$$

In section 9, we present the measured gravitational torques obtained with the copper SM. A Bayesian data analysis yields posterior distributions for G_{hom} in both the free-deflection and electrostatic-servo modes. It also provides the joint posterior distribution $p(\theta_0, G_a/G_{\text{hom}}|\text{data})$, shown in figure 14. The posterior mode gives $\theta_0 = 156°$ and $G_a/G_{\text{hom}} = 3.1 \times 10^{-5}$. The inferred amplitude agrees well with the numerical prediction, but the phase differs by nearly 120°, consistent with a possible misidentification of the fiducial-mark orientation which may have occurred when the masses were cleaned on arrival at NIST. In the posterior distribution, the calculated prediction is indicated by the black 'X'; to account for the phase offset, 120° was added to the theoretical θ_0.

The key conclusion of this section is that by performing a clocked rotation of all four SMs and averaging the results measured in the three rotational configurations, the effect of the density gradient can be canceled. Intuitively, averaging the three measurements is equivalent to a single measurement with the SMs effectively superposed in their three rotated orientations.

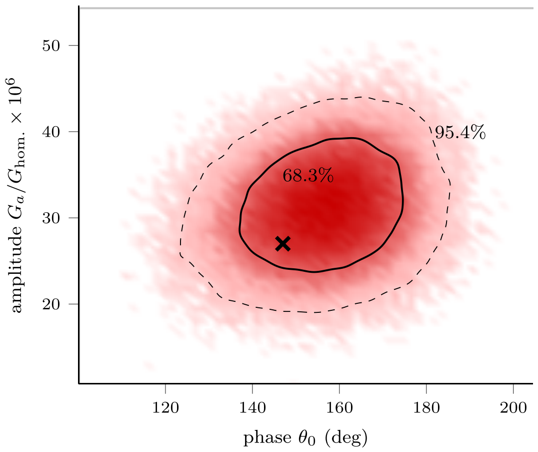

Figure 14. Posterior density in the $(\theta_0, G_a/G_{\text{hom}})$ plane, inferred from the copper source-mass measurements. The 68 % and 95 % credible regions (solid and dashed) reveal a well-constrained amplitude and a phase consistent with a 120° rotation ambiguity in the mass clocking. The black 'X' indicates the value expected from numerical integration of the measured density inhomogeneity.

Table 10. The mass integration produced by different simplifications of the system. The top line is the full mass integration and serves as the reference for the relative difference shown in the last column.

Description	Γ $/(\text{kg}^2\text{m}^{-1})$	Rel. diff. $\times 10^6$
Actual masses & disk, $\Delta\Gamma/2$	0.485 641	0
Actual masses & disk, $\Gamma_{\max}$	0.485 605	−73
Actual masses & no disk	0.485 099	−1117
Identical masses & no disk	0.485 099	−1116
Point masses & no disk	0.457 999	−56919

6.8. Uncertainty consideration for the mass integration

Table 10 summarizes the sensitivity of the mass integration. If we were to use only identical SMs and TMs without any other parts of the torsion balance, the calculated $\Gamma_{\max}$ would differ relatively by -1116×10^{-6} from $\Delta\Gamma/2$ obtained from the full calculation. Hence, if we want to trust the mass integration to within 11×10^{-6}, we have to make sure that the calculation of the extra parts and asymmetries are correct to 1 percent. We have the confidence that the mass integration is correct to that level. We assign a relative uncertainty of 11×10^{-6} to the mass integration.

In this section, we have mostly discussed the mass integration with the copper SMs at a particular MCR. The same logic applies to the mass integration for the sapphire SMs, but for that case the mass integration constant is only 44.8% of that of the copper SMs. The mass integration constants for the four experiments described in this article are given in table 15.

6.9. Calculation of the moment of inertia of the disk

The second numerical quantity required to determine G is the torsion disk's moment of inertia, I_{disk}, which enters only the free-deflection method. The electrostatic-servo method is static and independent of I_{disk}.

As shown in table 7, the disk contributes about 10% to the total moment of inertia but only 0.1% to $\Gamma_{\max}$. This disparity arises because, for a feature located at radius R_f, $I \propto R_\text{f}^2$ whereas $\Gamma_{\max} \propto R_\text{f}^4$. Thus, while $\Gamma_{\max}$ is dominated by a few features at large R_f, accurate modeling of I requires including the many complex features near the torsion strip to achieve a relative uncertainty of 10^{-5}.

The numerical integration includes all components listed in table 8. Assuming the non-test-mass parts are known to 1%, their 11% share of I would limit G_{free} to about 1×10^{-3} relative precision, motivating tighter verification.

To confirm the calculation, I_{disk} was (i) computed in SolidWorks, differing by only 1 g cm^2, and (ii) measured experimentally as described in section 8.4. The two methods agree within 1.6×10^{-4}, validating the mass-integration model and the underlying geometry data.

7. Thermal torque

Having established the mass-integration factor and moment-of-inertia model, we now turn to a thermal effect revealed during the NIST runs. As discussed in section 4, the vacuum pressure slowly rises after evacuation. This behavior ultimately revealed a significant systematic effect.

7.1. Empirical observations

During data analysis, the measured values of G were plotted against the residual-gas pressure (figure 15), revealing a strong correlation: higher pressures yielded larger values of G. Although the correlation was not perfect, it clearly indicated that the measurement depended on the vacuum pressure.

7.2. Deliberate pressure variation

In an initial test, the vacuum pressure was deliberately varied by partially closing the pumping-port valve. Figure 16 shows the measured torque. For low pressures, the torque depends linearly on the vacuum pressure p_v with a slope of 0.3×10^{-9} m^3.

At this point, it remains unclear how the pressure-dependent torque correlates with the SM position. Such a correlation is necessary to produce the consistent signal in the torque difference as shown in figure 16. To investigate the underlying mechanism, a dedicated heating test was performed.

7.3. Heating test

To probe a thermal origin, the SMs were removed and a heating tape was attached near TM 1 on the vacuum-vessel wall. The heater was placed a few degrees clockwise or counter-clockwise from the mass and modulated with a period of 24 min. The resulting torque signal (figure 17) closely mimicked the pressure-dependent response, supporting a thermally mediated effect. The fitted slope of 5.9×10^{-9} m^3 implies a temperature gradient about 30 times larger than during gravitational runs. The peak-to-peak temperature change on the vessel exterior was 55 mK (clockwise) and 22 mK (counter-clockwise), consistent with similar torque amplitudes despite the unequal external gradients.

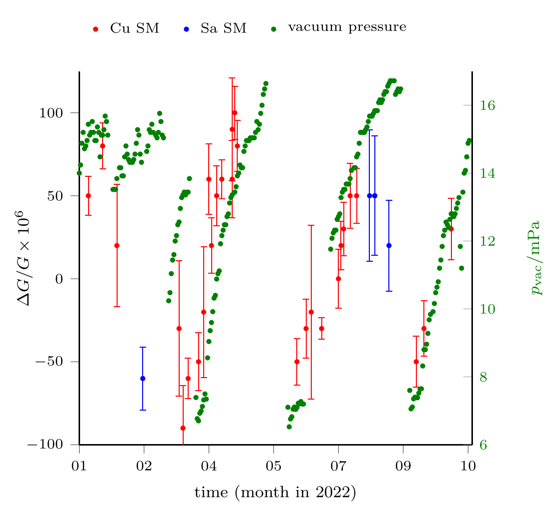

Figure 15. Correlation between measured G and residual-gas pressure. Red and blue points: G measurements (left axis); green points: chamber pressure (right axis).

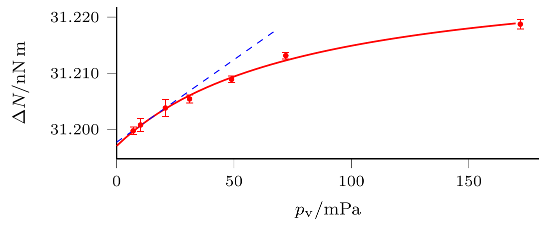

Figure 16. Intentionally changing the vacuum pressure alters the measured torque difference in free-deflection method. At low pressures (below approx. 25 mPa), the torque difference is well-approximated by a linear dependency on pressure, as indicated by the dashed line. This linear regime corresponds to the molecular flow limit, where the mean free path of the gas molecules exceeds the dimensions of the vacuum chamber (Knudsen number $Kn > 1$).

A temperature gradient along the vacuum wall can therefore generate a pressure-dependent torque. For this to appear in the G measurements, the gradient must vary with the SM position. One plausible cause is airflow: moving the SMs slightly changes the local heat transfer at the SM positions. Installing a simple cardboard shield atop the chamber to disrupt the laminar ceiling flow altered the measured torque by 8 μN m at 150 mPa.

The effect was therefore real and repeatable, but its underlying mechanism appears to require a kinetic-theory description. In the next section we

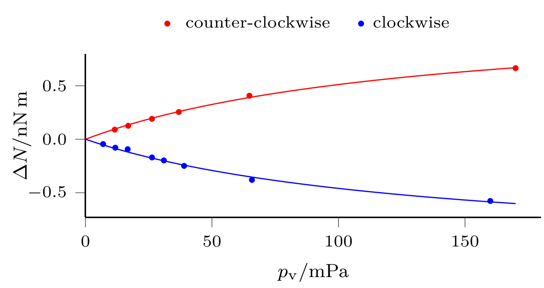

Figure 17. Torque difference measured with the free-deflection method while modulating the power of a heater placed near TM 1. The observed response is similar to that in figure 16, consistent with a thermal torque mediated by residual gas.

discuss the mediation of the spurious torque by momentum transfer from the residual gas.

7.4. Spurious torque from modulation of temperature gradients

For a gas in the transition regime between viscous and molecular flow, with a temperature gradient $\partial T/\partial x$, Waldmann [56] derived an expression for the thermophoretic force acting on a sphere of radius r_sph:

$$F_x = -\frac{8}{15} r_\mathrm{sph}^2 \lambda \sqrt{\frac{2\pi\, m_\mathrm{g}}{kT}}\, \frac{\partial T}{\partial x}. \tag{42}$$

Here, m_g is the molecular mass, λ the thermal conductivity and T the temperature of the gas. Using kinetic theory,

$$\lambda = \frac{f}{6} p_\mathrm{v} l \sqrt{\frac{8k}{T\pi\, m_\mathrm{g}}}, \tag{43}$$

where f is the number of degrees of freedom and l the mean free path. For an infinite volume, the mean free path depends on the vacuum pressure p_v, as

$$l_0 = \frac{kT}{\sqrt{2}\pi\, p_\mathrm{v} d_\mathrm{m}^2}, \tag{44}$$

where d_m is the molecular diameter. In a finite chamber, however, l is limited by geometry:

$$l = \left(\frac{1}{l_0} + \frac{1}{l_\mathrm{typ}}\right)^{-1}. \tag{45}$$

This interpolation bridges the molecular and viscous regimes.

Multiplying equation (42) by an effective lever arm r_eff gives the torque

$$N_\mathrm{gas} = -\frac{16f}{45} r_\mathrm{sph}^2 r_\mathrm{eff}\, p_\mathrm{v} \frac{l}{T} \frac{\partial T}{\partial x}, \tag{46}$$

where l is given by equation (45). All parameters (see table 11) are known or can be estimated except for l_typ and $\partial T/\partial x$, which were obtained from non-linear fits to the data in figures 16 and 17. The combined cross-sectional area of the four TMs was approximated as πr_sph^2.

Fits to the heating data gave $l_\mathrm{typ} = 0.124$ m and temperature gradients of 1.6 mK m^{-1} and -1.8 mK m^{-1} for the clockwise and counter-clockwise cases, respectively. For the gravitational data (figure 16), the best fit yielded $l_\mathrm{typ} = 0.209$ m and $\partial T/\partial x = 47.8$ µK m^{-1}. The fitted l_typ values agree with the chamber radius.

During the heating experiment, temperature changes on the chamber exterior were tens of mK, much larger than the internal gradients of interest. Because the relevant gradient acts inside the chamber, a large attenuation factor between the external and internal temperatures is expected. In the G measurements, a gradient of only 50 µKm^{-1} (a 5 µK difference across 10 cm) produces a thermally induced torque roughly 0.01% of the gravitational signal at the lowest pressures.

7.5. Temperature measurements

Figure 18 shows the vacuum-vessel geometry and sensor locations. Eight thermistors (Vishay 01M1002FF) were mounted on the upper horizontal step above the SMs, near the positions of maximum torque sensitivity.

A typical temperature record (figure 19) shows millikelvin-level variations synchronized with the SM motion. Some sensors warmed when the nearby mass approached, others when it receded, reflecting complex local airflow. A gradient of only 5 µK across 10 cm inside the can would suffice to produce the observed spurious torque—three orders of magnitude smaller than the external variations.

7.6. Comparison with the BIPM thermal environment

The BIPM apparatus operated in a larger, two-level laboratory with different thermal dynamics. The CMM and balance were housed in a 3 m × 4 m × 2.5 m cabin where the temperature remained within 0.1 K for periods of up to two weeks. We also note that the laboratory where the *Mark I* determination was performed was similar in size to that of the *Mark II*, but there was no cabin. The air circulation rate was significantly lower at the BIPM. There was no forced air flow in the cabin. Thus, the magnitude and even the sign of thermally induced torques could vary.

For comparison, at NIST the experiment was set up in a room with a footprint of 6.63 m × 3.5 m and a total height of 3.5 m. The vertical air speed in the room is estimated to be 0.2 m s^{-1}. Occasionally, a reversal of the correlation was occasionally observed: higher pressure led to a lower G. We hypothesize that this inversion resulted from altered airflow patterns that changed the direction of the temperature gradient inside the vacuum can. This is supported by the observation that even parking the CMM bridge in a

Table 11. Parameters used in the kinetic-gas torque calculation, assuming H_2O as the prevailing gas species.

Symbol	Description	Value	Unit
Constants and environment			
k	Boltzmann constant	1.38×10^{-23}	$\mathrm{J\,K^{-1}}$
T	Temperature	293.15	K
Gas properties (water vapor)			
m_g	Molecular mass	3.0×10^{-26}	kg
d_m	Molecular diameter	2.7×10^{-10}	m
f	Degrees of freedom	6	
Apparatus geometry			
r_{sph}	Test mass cross-sectional radius	5.6×10^{-2}	m
A_t	Area of 4 test masses	1.21×10^{-2}	$\mathrm{m^2}$
A_c	Chamber internal surface area	1.5	$\mathrm{m^2}$
r_{eff}	Effective lever arm	0.12	m
Experimental/derived			
l_{typ}	Typical dimension (from fit)	0.124	m
Calculated Prefactor			
$16 f r_{sph}^2 r_{eff} l_{typ}/(45T)$	$N_{gas}/(P_{vac}\partial T/\partial x)$	3.4×10^{-7}	$\mathrm{m^4\,K^{-1}}$

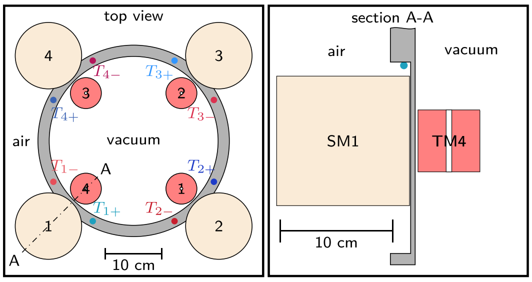

Figure 18. Vacuum-vessel layout and thermistor positions. Left: top view showing SMs, TMs, and eight thermistors (colored circles). Right: cross-section along AA, with the nearby thermistor T_{1+} projected below the overhang. The wall thickness is 4 mm near the masses, where the separation between the SM and wall is 1 mm.

different position could modify the flow enough to reverse the sign. Because such conditions cannot be reconstructed, no post hoc correction to earlier data is possible.

7.7. A path forward

The most direct mitigation is to lower the residual-gas pressure, since equation (46) predicts $N_{gas} \propto p_v$. Initially, the turbomolecular pump was mounted directly on the chamber via an ISO-K 100 flange, reducing the pressure by a factor of ten. Unexpectedly, the spurious torque increased: the 50 W heat load from the pump created warm airflow over the SMs, steepening the temperature gradient. Because $N_{gas} \propto p_v \partial T/\partial x$, the gain from lower pressure was canceled—or reversed—by the stronger gradient.

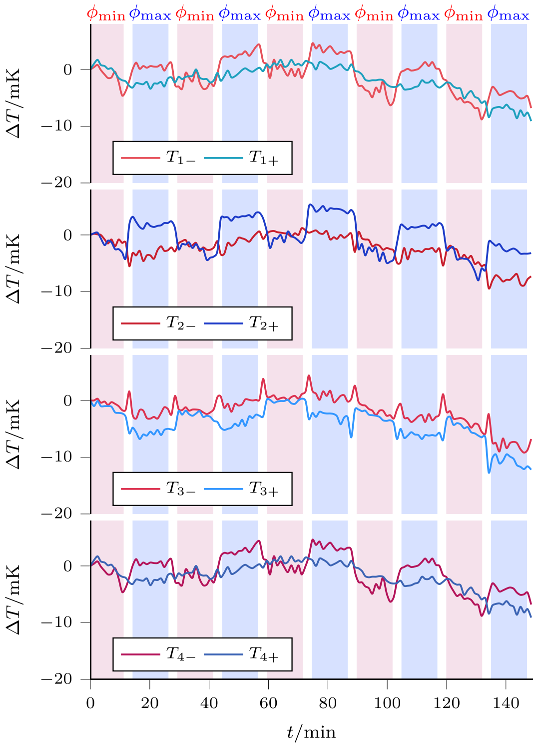

Figure 19. Relative temperature changes recorded by eight sensors around the vacuum vessel (see figure 18). Shaded areas indicate SM positions: blue for maximal clockwise torque, red for counter-clockwise. Some sensors (e.g. T_{2+}) show higher temperatures when the mass is nearby, others (e.g. T_{4-}) when it is away.

The final configuration placed the turbo pump laterally about 50 cm from the chamber and enclosed the entire apparatus in a Styrofoam box to block ceiling airflow. A slow air stirrer homogenized the internal air. A butterfly valve between the pump and vessel allowed operation at two pressures differing by about a factor of two (10 mPa and 20 mPa). Each G value was measured at both pressures and linearly extrapolated to $p_\mathrm{v} = 0$, yielding a relative uncertainty from the linear extrapolation of roughly 1.5×10^{-6}.

8. Auxiliary measurements

Having understood the thermal systematic, we now discuss the auxiliary measurements that constrain key systematics.

8.1. Measurements of the properties of the torsion strip

Anelasticity in flexure strips is well documented, notably by Quinn, Davis, and Speake [23–28, 5, 57, 58]. We summarize only what is needed for the free-deflection correction; the electrostatic-servo data are independent of κ and require no such correction.

The torsion pendulum obeys

$$I\ddot{\phi}_t + \beta\,\dot{\phi}_t + (\kappa_\mathrm{r} + i\kappa_\mathrm{i})\,\phi_\mathrm{t} = N_\mathrm{ext}, \tag{47}$$

with I the moment of inertia, β the gas-damping coefficient, and $\kappa = \kappa_\mathrm{r} + i\kappa_\mathrm{i}$ the complex torsion constant.

The strip is annealed Cu–Be (1.8% Be) with $l_\mathrm{st} = 160\,\mathrm{mm}$, $b_\mathrm{st} = 2.5\,\mathrm{mm}$, and $t_\mathrm{st} = 30\,\mu\mathrm{m}$. Twisting adds a lossless gravitational stiffness $\kappa_\mathrm{g} = Wb_\mathrm{st}^2/(12\,l_\mathrm{st})$ in parallel with the elastic stiffness $\kappa_\mathrm{e} = F_0\,b_\mathrm{st}t_\mathrm{st}^3/(3\,l_\mathrm{st})$, where $W = (m_\mathrm{disk} + 4\,m_\mathrm{t})\,g$ is the weight of the pendulum (g is the local acceleration measured to be $g = 9.80103143(4)\,\mathrm{m\,s^{-2}}$ at the nearby Kibble balance) and shear modulus $F_0 \approx 53\,\mathrm{GPa}$. Note that while κ_g represents a significant portion of the restoring torque, the total stiffness $\kappa = \kappa_\mathrm{e} + \kappa_\mathrm{g}$ is determined experimentally from the pendulum's period of oscillation and does not need to be calculated from these parameters with high accuracy. The real part of the torsion constant is

$$\kappa_\mathrm{r}(\omega) = \kappa_\mathrm{g} + \kappa_\mathrm{e}\left(1 + \frac{\delta F/F_0}{2\ln(\tau_\infty/\tau_0)}\ln\left(\frac{1+\omega^2\tau_\infty^2}{1+\omega^2\tau_0^2}\right)\right), \tag{48}$$

reflecting anelastic relaxation with $\tau \in [\tau_0, \tau_\infty]$ [26]. The imaginary part is modeled as

$$\kappa_\mathrm{i}(\omega) = \frac{\delta F}{3}\,\frac{b_\mathrm{st}t_\mathrm{st}^3}{l_\mathrm{st}}\,\frac{\arctan(\omega\tau_\infty) - \arctan(\omega\tau_0)}{\ln(\tau_\infty/\tau_0)}. \tag{49}$$

In the relevant band $\tau_\infty^{-1} \ll \omega \ll \tau_0^{-1}$,

$$\kappa_\mathrm{r} \approx \kappa_\mathrm{g} + \kappa_\mathrm{e}\left(1 + \frac{\xi}{\pi}\ln\left(\omega^2\tau_\infty^2\right)\right), \qquad \kappa_\mathrm{i} \approx \kappa_\mathrm{e}\,\xi, \tag{50}$$

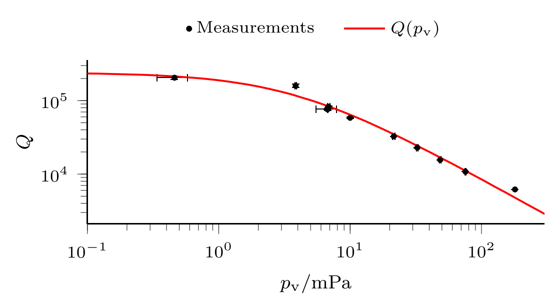

Figure 20. Quality factor Q versus residual gas pressure. Solid line: fit to equation (53).

with modulus defect

$$\xi = \frac{\pi\,(\delta F/F_0)}{2\ln(\tau_\infty/\tau_0)}. \tag{51}$$

At resonance ω_0,

$$\frac{1}{Q} = \frac{\kappa_\mathrm{i}}{\kappa_\mathrm{r}} + \frac{\beta\,\omega_0}{\kappa_\mathrm{r}} \approx \underbrace{\frac{\xi\,\kappa_\mathrm{e}}{\kappa_\mathrm{e}+\kappa_\mathrm{g}}}_{1/Q_\mathrm{max}} + \underbrace{\frac{\beta\,\omega_0}{\kappa_\mathrm{e}+\kappa_\mathrm{g}}}_{1/Q_\mathrm{gas}}. \tag{52}$$

In high vacuum $Q_\mathrm{gas} \propto 1/p_\mathrm{v}$ [59]. Writing $Q_\mathrm{gas} = p_Q/p_\mathrm{v}$, where p_Q is a constant to be determined, gives

$$Q(p) = \frac{p_Q\,Q_\mathrm{max}}{p_Q + p_\mathrm{v}\,Q_\mathrm{max}}. \tag{53}$$

During the pressure tests of section 7, we extracted Q from free-decay envelopes at multiple pressures (figure 20). Fitting equation (53) yields

$$Q_\mathrm{max} = 2.4(4) \times 10^5, \qquad p_Q = 8.8(6) \times 10^2\,\mathrm{Pa}. \tag{54}$$

These parameters calibrate $Q(p_\mathrm{v})$, enabling the pendulum to serve as an *in situ* pressure gauge; we use this in section 10 to infer the BIPM chamber pressure.

The most consequential result of the fit is Q_max, from which

$$\xi = \frac{\kappa_\mathrm{g}+\kappa_\mathrm{e}}{\kappa_\mathrm{e}}\,\frac{1}{Q_\mathrm{max}} = 1.18(17) \times 10^{-4}, \tag{55}$$

using the dilution factor $(\kappa_\mathrm{g}+\kappa_\mathrm{e})/\kappa_\mathrm{e} = 28.4$. This ξ agrees with literature [5, 60].

The effective stiffness depends weakly on the timescale. Using $\omega = 2\pi/T$,

$$\frac{\kappa_\mathrm{r}(T_0) - \kappa_\mathrm{r}(T_s)}{\kappa_\mathrm{g}+\kappa_\mathrm{e}} = \frac{\xi\,\kappa_\mathrm{e}}{\kappa_\mathrm{e}+\kappa_\mathrm{g}}\,\frac{2}{\pi}\ln\left(\frac{T_s}{T_0}\right). \tag{56}$$

For $T_s = 6T_0$ we obtain

$$\frac{\kappa_\mathrm{r}(T_0) - \kappa_\mathrm{r}(T_s)}{\kappa_\mathrm{g}+\kappa_\mathrm{e}} = 4.7(7) \times 10^{-6}. \tag{57}$$

Thus, κ_r at resonance is larger by 4.7(7) ppm than at the signal frequency. Since $G_\mathrm{free} \propto \kappa_\mathrm{r}(T_0) = \omega_0^2 I$ (section 3), we apply a relative correction of -4.7×10^{-6} with uncertainty 7×10^{-7} to G_free; this is included in its uncertainty budget.

Table 12. Comparison of measured and calculated torques for the empty torsion-disk configuration (TMs removed).

Source mass	Measured /pN m	Calculated /pN m	Difference /pN m
Copper	37.9(2)	38.1	0.2(2)
Sapphire	16.7(6)	17.4	−0.7(6)

8.2. Measurements with empty carousel

The SMs sit on an aluminum carousel: a hollow disk (outer diameter 570 mm, inner diameter 381 mm, thickness 30 mm). At each of the four stations, three blind holes form an equilateral triangle; ceramic balls in these holes kinematically engage V-grooves on the SMs for reproducible seating. With the carousel and balls installed but no SMs, we measured

$$N_{\mathrm{meas,no\,SM}} = -76\,\mathrm{fN\,m} \pm 169\,\mathrm{fN\,m}.$$

The calculation for an ideal, uniform-density aluminum disk yields

$$N_{\mathrm{calc,noSMs}} = 79\,\mathrm{fN\,m}. \tag{58}$$

The difference,

$$\Delta N_{\mathrm{no\,SM}} = 155\,\mathrm{fN\,m} \pm 169\,\mathrm{fN\,m},$$

is consistent with zero. We conservatively take 169 fN m (one sigma) as the background-torque uncertainty, i.e. relative fractions 5.5×10^{-6} (copper) and 1.1×10^{-5} (sapphire).

8.3. Measurements with empty torsion disk

Removing the TMs isolates the small contribution from the other components of the torsion balance, collectively referred to as disk, providing a check of the mass-integration model. As we have discussed in section 6.5, these components increase $\Gamma_{\max}$ relatively by 1.044×10^{-3}. After removing the TMs, we had to re-level the disk by redistributing balancing balls; the updated mass distribution was included in the model.

As shown in table 12, the copper configuration gave 37.9 pN m with 0.2 pN m (one-sigma) uncertainty; the measured-calculated difference 0.2(2) pN m is consistent with zero. A shorter sapphire run yielded −0.7(6) pN m, i.e. $\approx 1.2\sigma$.

We therefore assign 200 fN m (one sigma) to this part of the mass integration. The relative contributions of 6.5×10^{-6} (copper) and 1.3×10^{-5} (sapphire) are included as N_{zero} in the uncertainty budget.

8.4. Measurement of the moment of inertia

From equation (10), I_{disk} contributes about 11% to G_{free}, so its accuracy directly affects the combined uncertainty. Using mass and dimensional data (section 5) we compute (section 6.9)

$$I_{\mathrm{disk}} = 74403\,\mathrm{g\,cm}^2.$$

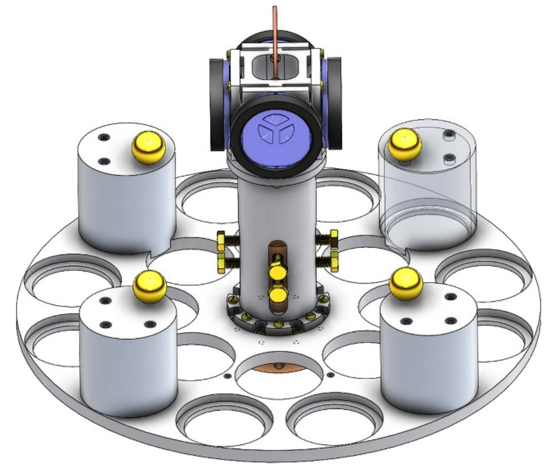

Figure 21. Technical drawing of the torsion pendulum with the aluminum 'cups' and brass balls installed. The cups replace the test masses during the moment-of-inertia measurement and serve both to enable electrostatic damping and to provide seating positions for the brass balls. By varying the ball positions while keeping the total suspended mass constant, the effective moment of inertia of the pendulum can be changed without altering the torsion constant of the fiber.

Table 13. Measured periods and calculated moments of inertia with cups installed and brass balls in two positions.

	Balls out	Balls in	Ratio, out/in
T/s	89.559(2)	87.261(6)	
T^2/s^2	8020.8(4)	7632.0(12)	1.05 095(16)
$I/\mathrm{g\,cm}^2$	153 681.7	146 327.1	1.05 094

To verify this value, we replaced the copper TMs with four thin-walled, upside-down aluminum cups ($m_{\mathrm{cup}} \approx 95.3\,\mathrm{g}$). These cups preserved the outer TM geometry, enabling the electrostatic damping system to function without modification. Each cup holds a brass ball ($m_{\mathrm{ball}} \approx 30.55\,\mathrm{g}$) that can be moved between 'in' and 'out' positions, altering the moment of inertia while keeping the total suspended mass constant. A drawing of the configuration is shown in figure 21.

With the balls in both positions, we measured the periods T_{in} and T_{out} and compared $T^2 \propto I$:

$$\frac{T_{\mathrm{out}}^2}{T_{\mathrm{in}}^2} = \frac{I_{\mathrm{out}}}{I_{\mathrm{in}}}.$$

The measured and calculated values (table 13) satisfy

$$\frac{T_{\mathrm{out}}^2 I_{\mathrm{in}}}{T_{\mathrm{in}}^2 I_{\mathrm{out}}} - 1 = (8 \pm 160)\times10^{-6},$$

fully consistent with unity. We therefore assign a conservative relative uncertainty of 1.6×10^{-4} to I_{disk}, in agreement with the numerical and SolidWorks analyses of section 6.9.

8.5. Calibration of the autocollimator at PTB

We measured ϕ_t with an Elcomat HR autocollimator (Möller-Wedel). The legacy controller occasionally produced outliers and was replaced.

A useful feature of our analysis is that the average of the free-deflection and servo results is, to first order, insensitive to a constant scale error in ϕ_t: $G_\text{free} \propto \phi_t$ while $G_\text{servo} \propto \phi_t^{-1}$, so the average, $G_\text{ave} = \frac{1}{2}(G_\text{free} + G_\text{servo})$ largely cancels a linear factor.

Let

$$G_\text{free} = c_f \phi_t, \quad G_\text{servo} = c_s \phi_t^{-1}.$$

Error propagation gives

$$\sigma_{G_\text{ave}} = \left| \frac{\partial G_\text{ave}}{\partial \phi_t} \right| \sigma_\phi = \frac{|G_\text{free} - G_\text{servo}|}{2} \frac{\sigma_\phi}{|\phi_t|},$$

which vanishes when $G_\text{free} = G_\text{servo}$. For the relative difference

$$\delta = \frac{G_\text{free} - G_\text{servo}}{G_\text{ave}},$$

one finds

$$\sigma_\delta = \left| \frac{\partial \delta}{\partial \phi_t} \right| \sigma_\phi = \frac{8 G_\text{free} G_\text{servo}}{(G_\text{free} + G_\text{servo})^2} \frac{\sigma_\phi}{|\phi_t|}.$$

The latter expression converges to $2\sigma_\phi/|\phi_t|$ for $G_\text{free} = G_\text{servo}$ Thus δ is proportional to the autocollimator uncertainty.

Non-linearity would break this cancellation, so the autocollimator was calibrated at PTB (Andreas Just) using established methods [61–65]. The PTB calibration file reports the deviation (true minus indicated angle) every $0.2''$ across the range $-60''$ to $60''$. Figure 22 shows these data together with the posterior mean of a Bayesian spline-augmented harmonic regression of the form

$$\mu(\phi) = B_\text{cos} \cos \omega\phi + B_\text{sin} \sin \omega\phi + S(\phi) \tag{59}$$

with $S(\phi)$ a cubic B-spline capturing slow variations in the autocollimator response. The shaded region includes the true curve with 68.4% posterior probability.

The frequency ω was restricted to the physically admissible range $[2\pi/16.32'', 2\pi/13.96'']$. To encode this range in the Bayesian model, we introduced an auxiliary variable $\nu \sim N(0.025, 0.03)$ and mapped it to ω using a scaled logistic transformation. The posterior mean of the resulting period is $15.448''$.

To assess the uncertainty in the harmonic component of the calibration model, we examined the joint posterior distribution of the sine and cosine amplitudes B_sin and B_cos. Figure 23 shows the corresponding posterior density together with the 68.3% and 95.4% highest-probability contours. The mean amplitude of the sinusoidal deviation $B_\text{amp} = \left(B_\text{cos}^2 + B_\text{sin}^2\right)^{1/2}$ is 4.63 mas. This amplitude determines the scale of the non-linear correction.

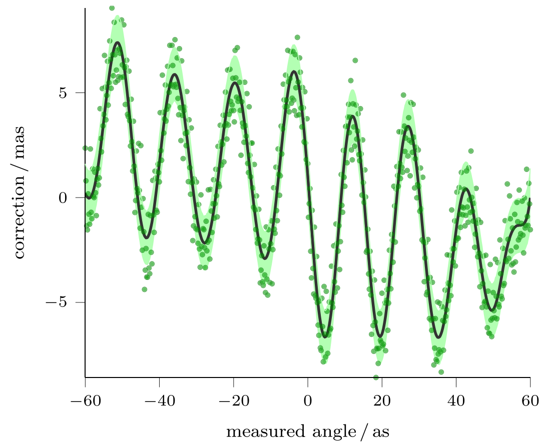

Figure 22. Autocollimator correction data (green points) measured at the PTB together with the posterior mean prediction of the regression model (solid line). The shaded region includes the true curve with 68.4% posterior probability.

The non-linearity in the autocollimator originates from both stray reflections and the discrete CCD structure. Because stray reflections can vary due to the optical path, the window and mirror from the experiment were sent to PTB for use during the calibration. Nevertheless, alignment shifts or spurious reflections introduce uncertainty in the phase of these non-linear features. To propagate this uncertainty, we generated an ensemble of random calibration curves with the same frequency and amplitude with phases distributed uniformly over the interval $[0, 2\pi]$. The bottom panel of figure 24 shows the distribution of the observed ϕ_t, and the top panel shows the resulting distribution of relative shifts in G with respect to the supplied calibration curve.

Because the non-linearity enters as a sinusoid with an unknown phase, each realization produces a shift $\Delta G/G = A \sin\phi$. Since ϕ is uniformly distributed, the resulting distribution of ΔG is the arcsine distribution with support $[a, b]$ and density

$$P(\Delta G/G) = \frac{1}{\pi\sqrt{(\Delta G/G - a)(b - \Delta G/G)}}, \tag{60}$$

For this distribution, the mean and variance are

$$E[\Delta G/G] = \frac{a+b}{2}, \quad \text{Var}(\Delta G/G) = \frac{(b-a)^2}{8}. \tag{61}$$

Since the G obtained from the PTB-supplied calibration curve lies close to one end of the arcsine interval, we treat the mean as a systematic shift and combine it in quadrature with the standard deviation to obtain the corresponding uncertainty. Table 14 summarizes

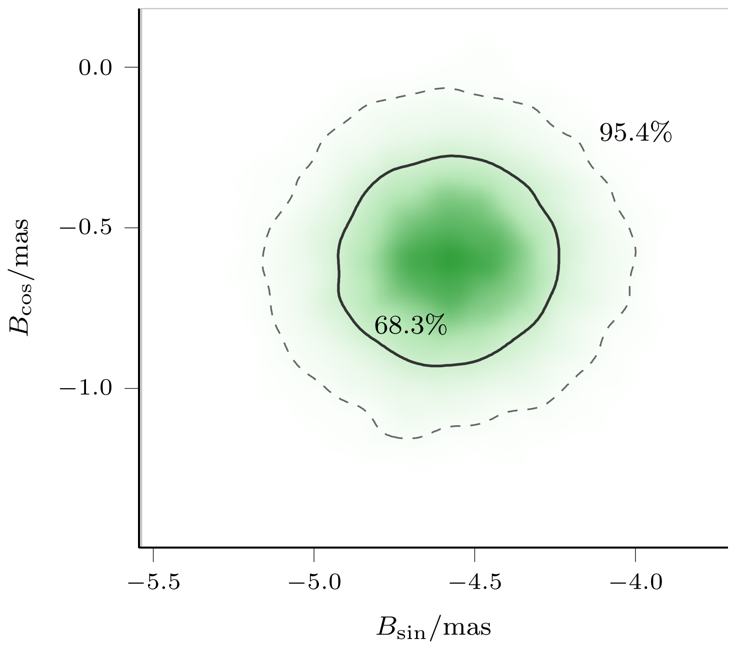

Figure 23. Joint Bayesian posterior for the harmonic coefficients $B_{\sin}$ and $B_{\cos}$ estimated from the spline-augmented harmonic regression. The 68.3% and 95.4% highest-posterior-density contours are shown.

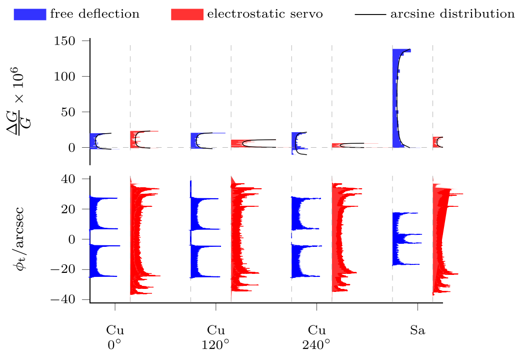

Figure 24. Effect of autocollimator non-linearity on the measured torsion angle and on the inferred value of G. Bottom: Histogram of the observed torsion angles ϕ_t for the four source mass configurations and the two measurement methods. Top: Histogram of the corresponding relative shifts in G with respect to the supplied calibration curve, obtained from the ensemble of randomized calibration curves. The sinusoidal non-linearity due to the CCD pixel structure has a fixed amplitude but an uncertain phase; each realization therefore produces a slightly different correction.

Table 14. Endpoints a and b of the support of the fitted arcsine distribution for the relative change ΔG due to autocollimator non-linearity, and the corresponding standard deviation σ. All values are relative to G.

Measurement	$a \times 10^6$	$b \times 10^6$	$u_r(\phi_t) \times 10^6$
Servo copper 0°	−1.3	23.7	14.2
Servo copper 120°	0.0	11.2	6.8
Servo copper 240°	−0.2	6.1	3.7
Free copper 0°	−2.8	20.4	12.0
Free copper 120°	−2.0	21.0	12.5
Free copper 240°	−10.7	22.1	12.9
Servo sapphire	−0.4	15.3	9.3
Free sapphire	−2.8	140.9	85.7

the fitted values of a and b together with resulting uncertainties.All quantities in the table are expressed relative to G.

For the four final results (copper servo, copper free, sapphire servo, and sapphire free) we assign the following relative uncertainties due to autocollimator non-linearity:

$$u_r(\phi_t) = (7.1,\ 12.5,\ 9.3,\ 85.7) \times 10^{-6} \tag{62}$$

For the copper results, the values in equation (62) represent the geometric averages of the relative uncertainties over the three, copper source-mass orientations listed in table 14, since the final copper values combine the results obtained in these configurations. The four uncertainties should be treated as fully correlated, since they arise from the same underlying autocollimator non-linearity model.

With all systematics characterized, we turn to the final determination of G.

9. Measurements of the gravitational constant

Having discussed the auxiliary measurements, we now focus on the measurements of the gravitational torque. The experimental data used to calculate the gravitational constant were collected between 24 January and 22 June 2024. Torque differences were measured using two methods: electrostatic-servo and free deflection. Measurements were taken with the

Table 15. Torque differences ΔN measured in the free-deflection and the electrostatic-servo modes for selected SM configurations. Preliminary G values follow from equation (1), $\Delta N = G\,16\,\Gamma_{\max}\,m_s m_t/R_s$. The values for m_s and m_t are listed in table 2. Uncertainties shown here are Type A only and for $k = 1$.

configuration	$\Delta N_{\text{free}}/\text{nN m}$	$\Delta N_{\text{servo}}/\text{nN m}$	R_s/mm	$\Gamma_{\max}$
Sapphire	13.9799 ± 0.0002	13.9778 ± 0.0003	213.9680	0.485 613
Copper 0°	31.1979 ± 0.0003	31.1962 ± 0.0004	213.9642	0.485 637
Copper 120°	31.1842 ± 0.0003	31.1828 ± 0.0006	213.9817	0.485 467
Copper 240°	31.1856 ± 0.0002	31.1836 ± 0.0003	213.9822	0.485 464

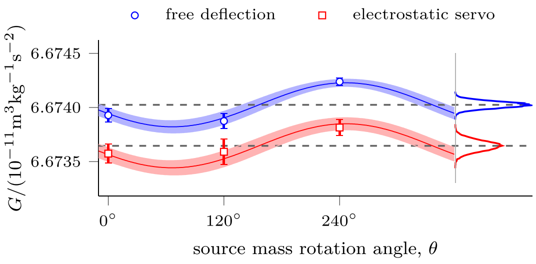

Figure 25. Six determinations of G obtained with copper source masses at three azimuthal rotation angles. A Bayesian analysis using the model $G(\theta) = G_i + A\sin(\theta - \theta_0)$, where $i \in \{1, 2\}$ corresponds to measurements performed with the electrostatic-servo (G_1) and free-deflection (G_2) methods, yields the marginal posterior distributions for G_1 and G_2 shown on the right. The joint posterior density in the (θ_0, G_a/G_{hom}) plane is presented in figure 14. The shaded bands surrounding the fits represent the 68.3% (1-sigma) posterior credible intervals.

sapphire SMs in one orientation and the copper SMs in three rotational orientations. The latter were necessary to average over a possible radial density gradient of each copper SM (see section 6.7), which was not present in the sapphire masses. All eight torque differences are listed in table 15. Immediately before or after each campaign, the MCR of the SMs was determined. The torsion balance remained under vacuum throughout, and the MCR of the TMs was measured only after all torque measurements were completed. The measured radii and the known SMs and TMs were used to calculate the mass-integration constants. These values are given in the last column of table 15. Dividing the measured torque differences by the mass-integration constant yields preliminary G values.

These six determinations of G measured with the copper masses were modeled using

$$G(\theta) = G_i + A\sin(\theta - \theta_0),$$

where $i \in \{1, 2\}$ indexes the two measurement methods. The Bayesian analysis yields the marginal posterior distributions for G_1 and G_2 shown in figure 25, and the joint posterior distribution of the parameters A and θ_0 is shown in figure 14. The posterior means and standard deviations were then combined with the corresponding quantities from the sapphire-mass measurements, giving four determinations of G labeled as Copper–servo, Copper–free, Sapphire–servo, and Sapphire–free. These values are listed in table 16.

Table 16. Four individual G determinations with their combined relative uncertainties ($k = 1$). Given the large number of observations underlying each determination, the statistical uncertainties are assumed to be characterized by effectively infinite degrees of freedom, and the uncertainty evaluations themselves are regarded as having negligible uncertainty.

Meas.	$G/(\text{m}^3\text{kg}^{-1}\text{s}^{-2})$	$u_r(G) \times 10^6$
Co. servo	6.673642×10^{-11}	23
Co. free	6.674021×10^{-11}	30
Sa. servo	6.672637×10^{-11}	38
Sa. free	6.673636×10^{-11}	94

Up to this point, we have considered only the Type-A uncertainty. We now turn to the evaluation of the combined uncertainty. The measurement result (value and uncertainty) for each G_i is determined by the values and uncertainties of 23 input quantities. The uncertainties of the input quantities are propagated to the G_i via known sensitivity coefficients using standard methods of uncertainty analysis. A summary of these uncertainties, grouped into 14 categories for each of the four measurements, is given in table 17. The four values of G, together with their combined standard uncertainties ($k = 1$), are listed in table 16.

Since several input variables enter into more than one measurement, the four resulting measured values of G are correlated. The corresponding correlation coefficients are listed in table 18. In contrast to the results of *Mark I* [4] and *Mark II* [5], all correlation coefficients are positive. This reflects the fact that the two dominant contributions to the uncertainty budget, the mass integration and the autocollimator non-linearity, enter with the same sign for all four determinations of G.

A detailed examination of the four determinations (see table 16) shows that, for both the copper and sapphire test-mass runs, the servo-mode values are consistently below the corresponding free-deflection values. It is noteworthy that the torque differences have similar magnitudes: 1.8 pN m ± 0.3 pN m for the copper run and 2.1 pN m ± 0.3 pN m for the sapphire run. The *Mark II* apparatus exhibited a comparable behavior with a difference of about 3.3 pN m.

Table 17. Relative uncertainty contributions sorted by size in the first column. All values are $k = 1$.

Item	$u_r(\text{item}) \times 10^6$ Copper–servo	$u_r(\text{item}) \times 10^6$ Copper–free	$u_r(\text{item}) \times 10^6$ Sapphire-servo	$u_r(\text{item}) \times 10^6$ Saphire–free
Mass integration	16.2	16.2	16.2	16.2
Autocoll non-linearity	10.7	14.2	9.3	85.7
Type A	7.6	4.6	19.5	13.9
Capacitance measurement	6.0	0.0	6.0	0.0
Voltages measurement	6.0	0.0	6.0	0.0
Zero torque	5.4	5.4	12.1	12.1
Pressure correction	1.6	1.6	21.5	21.5
SM MCR, R_t	1.0	1.0	1.0	1.0
Mass of one SM	0.2	0.2	0.4	0.4
Mass of one TM	0.1	0.0	0.1	0.0
MOI Disk, I_disk	0.0	19.8	0.0	19.8
TM MCR, R_t	0.0	2.9	0.0	2.9
Anelasticity	0.0	0.7	0.0	0.7
Period measurement	0.0	0.5	0.0	0.5
Combined	23.2	30.3	37.5	93.9

Table 18. Summary of relative uncertainties and correlations for the four G determinations. Bold diagonal entries are the relative standard uncertainties ($\times 10^6$). Off-diagonal entries are correlation coefficients.

	Copper servo	Copper free	Sapphire servo	Sapphire free
1	**23.2**	0.42	0.38	0.12
2		**30.3**	0.23	0.23
3			**37.5**	0.25
4				**93.9**

At present, we do not have a conclusive explanation for this systematic offset. One possibility is an additional thermal torque arising from resistive heating due to capacitive currents in the servo mode. To investigate this, we searched for a time-dependent torque following the application of the servo voltage; no significant time-dependent effect was observed. As of this writing, the origin of this behavior remains unknown. However, the fact that this discrepancy persists across both test-mass campaigns and was similarly observed in the *Mark II* apparatus suggests a reproducible, albeit unmodeled, physical bias.

We regard the observed offset as a source of overdispersion and address it by introducing 'dark uncertainty' [66] components that are taken into account as the four determinations $\{G_j\}$ are combined into a final consensus value while keeping the primary uncertainty budget intact. Rather than assuming that the same amount of dark uncertainty affects the four determinations equally, the Bayesian hierarchical model that we use to produce the consensus value explicitly entertains the possibility of there being configuration-specific, yet unidentified biases.

The measurement model fitted to the results is

$$G_j = \mu + \lambda_j + \varepsilon_j \tag{63}$$

for the four configurations $j \in \{1,2,3,4\}$ where μ denotes the consensus value of G. The term λ_j represents a configuration-specific effect, reflecting whether a specific experimental setup tends to yield a systematically high or low estimate. The term ε_j represents the measurement error incurred in that configuration, whose typical size is evaluated in the corresponding uncertainty budget.

The configuration effects $\{\lambda_j\}$ are assumed centered at zero, with standard deviations $\{\tau_j\}$. These biases become apparent when the four estimates of G are intercompared. They are modeled by a zero-mean multivariate normal distribution, $\boldsymbol{\lambda} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where the covariance matrix $\boldsymbol{\Sigma}$ is constructed by scaling the correlation matrix $\mathbf{R}$ with the configuration-specific standard deviations τ_j such that $\Sigma_{ij} = \tau_i \tau_j R_{ij}$.

The model was calibrated using a Bayesian procedure [67–69]. For the prior distribution for μ we use a Gaussian centered at the CODATA recommended value $G_\text{CODATA} = 6.6743 \times 10^{-11}\,\text{m}^3\text{kg}^{-1}\text{s}^{-2}$ [19]. For the standard deviation we use $9 \times 10^{-15}\,\text{m}^3\text{kg}^{-1}\text{s}^{-2}$ which corresponds to the robust estimate (Q_n) [70] of the standard deviation of the data points shown in figure 1. The four components of τ are assigned independent half-Cauchy prior distributions with medians scaled by a reference value $\tau_G = 9.6 \times 10^{-15}\,\text{m}^3\text{kg}^{-1}\text{s}^{-2}$ [71]. These medians reflect an expert ranking of configuration reliability: τ_G for the copper-free measurement (the most reliable), $2\tau_G$ for copper-servo, $3\tau_G$ for sapphire-free, and $4\tau_G$ for sapphire-servo.

The consensus estimate μ and its uncertainty were computed using robust estimators (Huber's M [72] and Q_n, respectively) applied to large samples drawn from the posterior distribution of the consensus value, μ. The estimates for τ_j given in table 19 are the medians of their posterior distributions. The

Table 19. Bayesian consensus results. The table lists the configuration-specific dark uncertainty (τ_j) and the estimated configuration effect (λ_j) for each measurement mode. The Copper-Free result carries the most weight due to its lower dark uncertainty.

Measurement	Dark uncertainty $\tau_j/\mathrm{m^3\,kg^{-1}\,s^{-2}}$	Config. Effect $\lambda_j/\mathrm{m^3\,kg^{-1}\,s^{-2}}$
Copper servo	7.24×10^{-15}	-2.03×10^{-15}
Copper free	5.08×10^{-15}	9.76×10^{-16}
Sapphire servo	2.20×10^{-14}	-1.20×10^{-14}
Sapphire free	1.02×10^{-14}	-1.57×10^{-15}

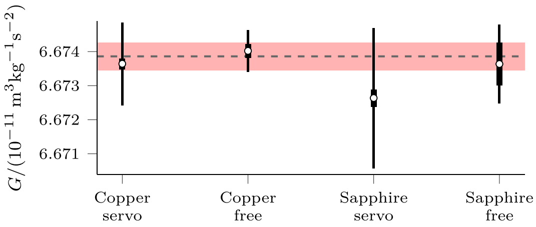

Figure 26. Comparison of the four determinations of the gravitational constant. Thick error bars represent the reported experimental uncertainties, while thin error bars represent the total uncertainty, calculated by combining experimental and 'dark' uncertainties in quadrature (root-sum-square). The dashed line indicates the Bayesian consensus estimate of G, with the shaded region representing its standard uncertainty ($k = 1$).

estimates of the $\{\lambda_j\}$ are Huber's M estimates derived from large samples drawn from the corresponding posterior distributions.

Our final combined result for the determination of the gravitational constant is

$$G = (6.67387 \pm 0.00038) \times 10^{-11}\,\mathrm{m^3 kg^{-1} s^{-2}}. \tag{64}$$

This corresponds to a relative standard uncertainty of 5.7×10^{-5}.

The analysis also quantifies the dark uncertainty specific to each configuration. As shown in table 19 the posterior analysis assigns the lowest configuration-specific dark uncertainty to the copper-free and copper-servo measurements, while identifying larger unexplained effects in the sapphire measurements. This statistical inference aligns with our experimental intuition that the sapphire measurements were more susceptible to unmodeled systematics, potentially related to the lower signal-to-noise ratio or electrostatic variations.

The posterior means of the configuration effects λ_j indicate that the copper-free measurement lies closest to the consensus value ($\lambda_2 \approx 9.76 \times 10^{-16}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$), whereas the Sapphire-Servo method exhibits the largest deviation ($\lambda_3 \approx -1.20 \times 10^{-14}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$). A visual summary of these four individual determinations, their combined uncertainties including the dark uncertainty contributions, and the resulting consensus estimate is presented in figure 26.

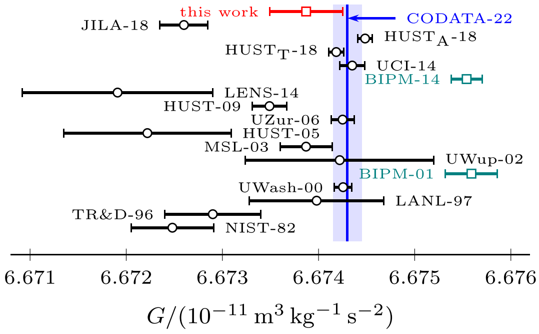

Figure 27. Comparison of this work with selected recent measurements. All uncertainty bars correspond to a coverage factor $k = 1$ (1-sigma uncertainty). The solid vertical line and blue shaded area indicate the 2022 CODATA [19] recommended value and its stated uncertainty. Teal squares denote measurements made with the BIPM torsion balance.

Figure 27 compares our result with previous measurements and with the CODATA recommended value of the gravitational constant. Our result differs relatively from the CODATA value by -6.4×10^{-5} and is inconsistent with the earlier determination of G obtained using the same apparatus. This discrepancy is examined in the following section.

10. Parting words

The preceding analysis defines our best estimate of G; in closing, we reflect on what this replication teaches about precision and reproducibility. Nearly a decade ago we set out to replicate a measurement of the gravitational constant with the aim of sharpening the global determination of G. While at first glance this result may not appear to sharpen the global picture of G, we believe the rigorous techniques developed here add valuable tools and clarity to the field.

A blind analysis was implemented to prevent experimenter bias. NIST's Mass and Force Group multiplied each SM value by an unknown factor $(1 + r)$ with $r \in [-1 \times 10^{-3}, +1 \times 10^{-3}]$, stored in a secure envelope. The envelope was opened during a public talk at CPEM 2024 (Colorado), once we were satisfied with the internal consistency. The G value reported at CPEM differed by -4.4×10^{-5} from the value reported here; small subsequent changes arose from bookend measurements of the MCR of the TMs.

Our result does not corroborate either the earlier BIPM value or the current CODATA recommendation. Relative to CODATA, the BIPM determination lies at $+1.71 \times 10^{-4}$ and ours at -6.4×10^{-5}, placing the two on opposite sides. However, the difference between the value we measured for G and the CODATA recommended value does not differ significantly from zero (p-value $= 0.29$) when the respective uncertainties are taken into account.

While we did not identify a single dominant cause for the discrepancy with the BIPM result, we were able to quantify a number of systematic effects not previously treated for this apparatus, including: (1) the trilobed shape of the original TMs, which can bias R_t (section 5); (2) gas-pressure forces producing a spurious torque during SM motion (section 7); (3) the non-linearity of the autocollimator (section 8.5); and (4) the dark uncertainty implied by the discrepancy between our individual measurements (section 9).

If R_t were underestimated by 1 µm ($\Delta R_t/R_t \approx 8.3 \times 10^{-6}$), the relative sensitivities (−2.4 free, −4.4 servo) would lower G_free by $\sim 20 \times 10^{-6}$ and G_servo by $\sim 37 \times 10^{-6}$. Quinn *et al* [5] report $G_\text{servo} - G_\text{free} = -104 \times 10^{-6}$. This correction would shift the difference by $\sim -17 \times 10^{-6}$, increasing the discrepancy. According to table 3, a difference of 1 µm between full and partial coverage is geometrically possible. However, the TM at NIST differed from those at the BIPM. Our measurements show that sampling only a single partial circle at mid-height could bias the observed radius by up to 2.5 µm. While this effect could theoretically explain a fraction of the G shift, it creates a contradiction: the measured difference between free-deflection and servo-mode results observed at the BIPM would increase.

For R_s, switching from partial to full CMM coverage changes the estimate for the MCR of the SMs by ~ 500 nm ($\Delta R_s/R_s \approx 2.3 \times 10^{-6}$). With $c_G(R_s) \approx 5.4$, both G_free and G_servo would shift by only $\approx 12.6 \times 10^{-6}$, too small to be significant.

No archived record exists of the vacuum pressure at the BIPM during the time the experiment was originally performed. Quinn *et al* [5] reports a pendulum quality factor Q between 1×10^5 and 3×10^5. Using equation (53) and figure 20, this implies $p \lesssim 1$ mPa. At NIST we observed $\partial(\Delta N)/\partial p = 3 \times 10^{-10}\,\text{m}^3$ (figure 16), so $p \leq 1$ mPa would yield $|\Delta N| \leqslant 0.3$ pN m, i.e. $|\Delta G|/G \approx 1 \times 10^{-5}$. This torque is too small to explain the discrepancy; however, thermal torques may become significant in experiments that employ smaller signal torques.

Collectively, these systematic checks constrain the potential impact of such effects to the 3×10^{-5} level, leaving the larger discrepancy unresolved. Nevertheless, the lesson is clear: standard assumptions about geometric perfection and vacuum levels are no longer sufficient at this level of precision. For future efforts, we therefore recommend: (1) moving beyond partial metrology to routine full-coverage CMM mapping; and (2) actively modeling gas-pressure torques arising from even minor temperature gradients.

Finally, we note that our total uncertainty is larger than in the previous determination. This increase does not reflect a degradation in measurement performance, but rather a more complete accounting of systematic limits. By explicitly including the 'dark uncertainty' [66] driven by the scatter among our individual datasets, we believe this result provides a more robust estimate of the true accuracy achievable with this apparatus.

In closing, while the historical plot of G values may appear messier with the inclusion of this result, we hope to have added clarity in approach, technique, and craftsmanship to the field. Precision metrology is not merely about converging on a number; it is about the rigorous exposure of unknowns. By presenting these data, imperfect but honestly characterized, we aim to provide a solid, if complex, footing for future efforts to establish a trustworthy estimate of G.

Acknowledgments

This measurement benefited from the contributions of many colleagues. We thank Ralf Geckeler and Andreas Just (PTB) for calibrating the autocollimator. We also thank our NIST colleagues Richard Steiner, Andrew Koffman, and Patrick Abbott for calibrating the ac voltmeter, the capacitance bridge, and the masses, respectively. Finally, this experiment would not have been possible without the support of Carl Williams and Terry Quinn, whose efforts in organizing the G workshop at NIST and facilitating the transfer of the apparatus were essential.

The authors thank our NIST colleague Jon Pratt and our former NIST colleague Richard Davis for proofreading the manuscript. We also thank the two anonymous reviewers for their careful work and comments that improved the manuscript.

ORCID iDs

Stephan Schlamminger 0000-0002-9270-4018
Leon Chao 0000-0001-7589-4019
Vincent Lee 0000-0002-2953-6322
Craig Shakarji 0009-0000-9877-7034
Antonio Possolo 0000-0002-8691-4190
David Newell 0000-0002-2612-1172
Julian Stirling 0000-0002-8270-9237
Clive Speake 0000-0002-2031-7449

References

[1] Luther G G and Towler W R 1982 Redetermination of the Newtonian gravitational constant g *Phys. Rev. Lett.* **48** 121
[2] Karagioz O V and Izmailov V P 1996 Measurement of the gravitational constant with a torsion balance *Meas. Tech.* **39** 979
[3] Gundlach J H and Merkowitz S M 2000 Measurement of Newton's constant using a torsion balance with angular acceleration feedback *Phys. Rev. Lett.* **85** 2869–72
[4] Quinn T J, Speake C C, Richman S J, Davis R S and Picard A 2001 A new determination of G using two methods *Phys. Rev. Lett.* **87** 111101
[5] Quinn T J, Speake C, Parks H and Davis R 2014 The BIPM measurements of the Newtonian constant of gravitation, G *Phil. Trans. R. Soc. A* **372** 20140032

[6] Bagley C H and Luther G G 1997 Preliminary results of a determination of the Newtonian constant of gravitation: a test of the kuroda hypothesis *Phys. Rev. Lett.* **78** 3047

[7] Kleinevoß U 2002 Bestimmung der newtonschen gravitationskonstanten *G PhD Thesis* (University of Wuppertal)

[8] Armstrong T R and Fitzgerald M P 2003 New measurements of *g* using the measurement standards laboratory torsion balance *Phys. Rev. Lett.* **91** 201101

[9] Luo J, Hu Z-K, Fu X-H, Fan S-H and Tang M-X 1998 Determination of the Newtonian gravitational constant G with a nonlinear fitting method *Phys. Rev. D* **59** 042001

[10] Hu Z-K, Guo J-Q and Luo J 2005 Correction of source mass effects in the HUST-99 measurement of *G Phys. Rev. D* **71** 127505

[11] Schlamminger S, Holzschuh E, Kündig W, Nolting F, Pixley R E, Schurr J and Straumann U 2006 Measurement of Newton's gravitational constant *Phys. Rev. D* **74** 082001

[12] Luo J, Liu Q, Tu L-C, Shao C-G, Liu L-X, Yang S-Q, Li Q and Zhang Y-T 2009 Determination of the Newtonian gravitational constant *G* with time-of-swing method *Phys. Rev. Lett.* **102** 240801

[13] Tu L-C, Li Q, Wang Q-L, Shao C-G, Yang S-Q, Liu L-X, Liu Q and Luo J 2010 New determination of the gravitational constant *G* with time-of-swing method *Phys. Rev. D* **82** 022001

[14] Parks H V and Faller J E 2010 Simple pendulum determination of the gravitational constant *Phys. Rev. Lett.* **105** 110801

[15] Parks H V and Faller J E 2014 A simple pendulum laser interferometer for determining the gravitational constant *Phil. Trans. R. Soc. A* **372** 20140024

[16] Rosi G, Sorrentino F, Cacciapuoti L, Prevedelli M and Tino G M 2014 Precision measurement of the Newtonian gravitational constant using cold atoms *Nature* **510** 518

[17] Newman R, Bantel M, Berg E and Cross W 2014 A measurement of *G* with a cryogenic torsion pendulum *Phil. Trans. R. Soc. A* **372** 20140025

[18] Li Q *et al* 2018 Measurements of the gravitational constant using two independent methods *Nature* **560** 582–8

[19] Mohr P J, Newell D B, Taylor B N and Tiesinga E 2025 CODATA recommended values of the fundamental physical constants: 2022 *Rev. Mod. Phys.* **97** 025002

[20] National Institute of Standards and Technology 2016 NSF ideas laboratory @ NIST: measuring "Big G" challenge (available at: www.nist.gov/news-events/events/nsf-ideas-laboratory-nist-measuring-big-g-challenge) (Accessed 02 June 2025)

[21] Osborne J W and Overbay A 2004 The power of outliers (and why researchers should always check for them) *Pract. Assess. Res. Eval.* **9** 6

[22] National Institute of Standards and Technology 2014 The trouble with capital 'G (available at: www.nist.gov/news-events/news/2014/10/trouble-capital-g) (Accessed 02 June 2025)

[23] Quinn T J, Speake C C, Brown L M 1992 Materials problems in the construction of long-period pendulums *Phil. Mag.* **65** 261–76

[24] Quinn T J, Speake C C, Davis R S and Tew W 1995 Stress-dependent damping in CuBe torsion and flexure suspensions at stresses up to 1.1 gpa *Phys. Lett. A* **197** 197–208

[25] Quinn T J, Davis R S, Speake C C and Brown L M 1997 The restoring torque and damping in wide CuBe torsion strips *Phys. Lett. A* **228** 36–42

[26] Speake C C, Quinn T J, Davis R S and Richman S J 1999 Experiment and theory in anelasticity *Meas. Sci. Technol.* **10** 430

[27] Quinn T J, Speake C C and Davis R S 1986 A 1 kg mass comparator using flexure-strip suspensions: preliminary results *Metrologia* **23** 87

[28] Quinn T J 1992 The beam balance as an instrument for very precise weighing *Meas. Sci. Technol.* **3** 141–59

[29] Quinn T J 1991 The kilogram: the present state of our knowledge *IEEE Trans. Instrum. Meas.* **40** 81–85

[30] Speake C C and Quinn T J 1989 A search for composition-dependent gravity using a beam balance *IEEE Trans. Instrum. Meas.* **38** 189–95

[31] Kuroda K 1995 Does the time-of-swing method give a correct value of the Newtonian gravitational constant? *Phys. Rev. Lett.* **75** 2796–8

[32] de Boer H, Haars H, Michaelis W and Schlimme E 1980 Quadrantenelektrometer als Drehmomentmesser für kleine Drehmomente *Feinw. Tech. Messtech.* **88** 237–41

[33] de Boer H, Haars H, Michaelis W and Schlimme E 1987 A new experiment for the determination of the Newtonian gravitational constant *Metrologia* **24** 171

[34] Michaelis W, Haars H and Augustin R 1995 A new precise determination of Newton's gravitational constant *Metrologia* **32** 267

[35] Cohen E R and Taylor B N 1987 The 1986 adjustment of the fundamental physical constants *Rev. Mod. Phys.* **59** 1121–48

[36] Michaelis W, Melcher J and Haars H 2004 Supplementary investigations to PTB's evaluation of *G Metrologia* **41** L29

[37] Speake C C, Davis R S, Quinn T J and Richman S J 1999 Electrostatic damping and its effect on precision mechanical experiments *Phys. Lett. A* **263** 219–25

[38] Quinn T J, Speake C C and Davis R S 1997 Novel torsion balance for the measurement of the newtonian gravitational constant *Metrologia* **34** 245

[39] Boys C V 1895 I. On the Newtonian constant of gravitation *Phil. Trans. R. Soc. London* **186** 1–72

[40] Engelsen N J, Beccari A and Kippenberg T J 2024 Ultrahigh-quality-factor micro- and nanomechanical resonators using dissipation dilution *Nat. Nanotechnol.* **19** 725–37

[41] Richman S J, Quinn T J, Speake C C and Davis R S 1999 Preliminary determination of G using the BIPM torsion strip balance *Meas. Sci. Technol.* **10** 460

[42] Schlamminger S, Gundlach J H and Newman R D 2015 Recent measurements of the gravitational constant as a function of time *Phys. Rev. D* **91** 121101

[43] Quinn T, Parks H, Speake C and Davis R 2013 Improved determination of *G* using two methods *Phys. Rev. Lett.* **111** 101102
Quinn T, Parks H, Speake C and Davis R 2014 *Phys. Rev. Lett.* **113** 039901 (erratum)

[44] Quinn T, Speake C, Parks H and Davis R 2014 Erratum: Improved determination of *G* using two methods [phys. rev. lett. 111, 101102 (2013)] *Phys. Rev. Lett.* **113** 039901

[45] Quinn T 2014 Outcome of the royal society meeting on G held at Chicheley Hall on 27 and 28 February 2014 to discuss 'the Newtonian constant of gravitation, a constant too difficult to measure? *Phil. Trans. R. Soc. A* **372** 20140286

[46] Schlamminger S, Chao L, Lee V, Newell D and Speake C 2022 The crane operator's tricks and other shenanigans with a pendulum *Am. J. Phys.* **90** 169

[47] Speake C C 2005 Newton's constant and the twenty-first century laboratory *Phil. Trans. R. Soc. A* **363** 2265–87

[48] Smythe W R 1939 *Static and Dynamic Electricity* (McGraw-Hill)

[49] Schlamminger S, Chao L, Lee V, Newell D B and Speake C C 2022 Design of electrostatic feedback for an experiment to measure *G IEEE Open J. Instrum. Meas.* **1** 19

[50] Kibble B P 2014 G-up (or down?) [basic metrology] *IEEE Instrum. Meas. Mag.* **17** 16–17

[51] Speake C C, Bryant J L, Davis R S and Quinn T J 2023 The influence of time varying magnetic fields on the determinations of Newton's constant at the Bureau International des Poids et Mesures *Metrologia* **60** 024001

[52] Gapinski B and Wieczorowski M 2014 Measurement of diameter and roundness on incomplete outline of element with three-lobbing deviation *Proc. Eng.* **69** 247–54

[53] Stoup J R and Doiron T D 2001 Accuracy and versatility of the NIST M48 coordinate measuring machine *Recent*

Developments in Traceable Dimensional Measurements vol 4401, ed J E Decker and N Brown (International Society for Optics and Photonics) pp 136–46

[54] Stirling J and Schlamminger S 2019 Closed form expressions for gravitational multipole moments of elementary solids *Phys. Rev. D* **100** 124053

[55] Speake C C 2025 Source mass homogeneity *Internal Report* (Unpublished)

[56] Waldmann L 1959 Über die Kraft eines inhomogenen Gases auf kleine suspendierte Kugeln *Z. Naturfor. A* **14** 589

[57] Speake C C 1987 Fundamental limits to mass comparison by means of a beam balance *Proc. R. Soc. A* **414** 333–58

[58] Speake C C 2018 Anelasticity in flexure strips revisited *Metrologia* **55** 114

[59] Braginsky V B and Manukin A B 1977 *Measurement of Weak Forces in Physics Experiments* (University of Chicago Press)

[60] Keck L, Schlamminger S, Theska R, Seifert F and Haddad D 2024 Flexures for kibble balances: minimizing the effects of anelastic relaxation *Metrologia* **61** 045006

[61] Probst R, Wittekopf R, Krause M, Dangschat H and Ernst A 1998 The new PTB angle comparator *Meas. Sci. Technol.* **9** 1059

[62] Just A, Krause M, Probst R and Wittekopf R 2003 Calibration of high-resolution electronic autocollimators against an angle comparator *Metrologia* **40** 288

[63] Geckeler R D, Just A, Krause M and Yashchuk V V 2010 Autocollimators for deflectometry: Current status and future progress *Nucl. Instrum. Methods Phys. Res. A* **616** 140–6

[64] Geckeler R D, Link A, Krause M and Elster C 2014 Capabilities and limitations of the self-calibration of angle encoders *Meas. Sci. Technol.* **25** 055003

[65] Geckeler R D, Krause M, Just A, Kranz O and Bosse H 2015 New frontiers in angle metrology at the PTB *Measurement* **73** 231–8

[66] Thompson M and Ellison S L R 2011 Dark uncertainty *Accr. Qual. Assur.* **16** 483–7

[67] Meija J, Bodnar O and Possolo A 2023 Ode to Bayesian methods in metrology *Metrologia* **60** 052001

[68] Rinaldi S, Middleton H, Pozzo W D and Gair J 2023 Bayesian analysis of systematic errors in the determination of the constant of gravitation *Eur. Phys. J. C* **83** 891

[69] Dose V 2006 Bayesian estimate of the Newtonian constant of gravitation *Meas. Sci. Technol.* **18** 176

[70] Rousseeuw P J and Croux C 1993 Alternatives to the median absolute deviation *J. Am. Stat. Assoc.* **88** 1273–83

[71] Possolo A 2023 Tracking truth through measurement and the spyglass of statistics *Stat. Sci.* **38** 655–71

[72] Huber P J and Ronchetti E M 2009 *Robust Statistics* 2nd edn (Wiley)