MICROSCOPE Mission: Final Results of the Test of the Equivalence Principle

Pierre Touboul, Gilles Métris, Manuel Rodrigues, Joel Bergé, Alain Robert, Quentin Baghi, Yves André, Judicaël Bedouet, Damien Boulanger, Stefanie Bremer, et al.

To cite this version:

HAL Id: hal-03781337
https://hal.archives-ouvertes.fr/hal-03781337
Submitted on 21 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
MIRCOSCOPE Mission: Final Results of the Test of the Equivalence Principle

Pierre Touboul,1,* Gilles Métris,2,† Manuel Rodrigues,‡,4 Joel Bergé,3 Alain Robert,4 Quentin Baghi,2,3,8 Yves André,4 Judicaël Bedouet,5 Damien Boulanger,3 Stefanie Bremer,6,1 Patrice Carle,1 Ratana Chhun,3 Bruno Christophe,3 Valerio Cipolla,4 Thibault Damour,2 Pascale Danto,4 Louis Demange,2 Hansjoerg Dittus,8 Océane Dhuique,3 Pierre Fayet,9 Bernard Foulon,3 Pierre-Yves Guidotti,4,5 Daniel Hagedorn,10 Emilie Hardy,3 Phuong-Anh Huynh,3 Patrick Kayser,2 Stéphanie Lala,1 Claus Lämmerzahl,6 Vincent Lebat,3 Françoise Liorzou,3 Meike List,6,† Frank Löffler,10 Isabelle Panet,1 Martin Pernot-Borràs,7 Laurent Perraud,4 Sandrine Pires,12 Benjamin Pouilloux,4,* Pascal Prieur,4 Alexandre Rebray,3 Serge Reynaud,1 Benny Rievers,6 Hanns Selig,6,† Laura Serron,2 Timothy Sumner,14 Nicolas Tanguy,3 Patrizia Torresi,4 and Pieter Visser15

1ONERA, Université Paris Saclay, F-91123 Palaiseau, France
2Université Côte d’Azur, Observatoire de la Côte d’Azur, CNRS,IRD, Géoaizur, 250 avenue Albert Einstein, F-06560 Valbonne, France
3DPHY, ONERA, Université Paris Saclay, F-92322 Châtillon, France
4CNES Toulouse, 18 avenue Edouard Belin—31401 Toulouse Cedex 9, France
5ONERA, Université de Toulouse, F-31055 Toulouse, France
6ZARM, Center of Applied Space Technology and Microgravity, University of Bremen, Am Fallturm, D-28359 Bremen, Germany
7IHES, Institut des Hautes Études Scientifiques, 35 Route de Chartres, 91440 Bures-sur-Yvette, France
8DLR, Köln headquarters, Linder Höhe, 51147 Köln, Germany
9Laboratoire dephysique de l’École normale supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université de Paris, F-75005 Paris, France, and CPHT, École polytechnique, IPP, F-91128 Palaiseau, France
10PTB, Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
11IPGP, 35 rue Hélène Brion, 75013 Paris, France
12Université Paris Saclay et Université de Paris, CEA, CNRS, AIM, F-91190 Gif-sur-Yvette, France
13Laboratoire Kastler Brossel, Sorbonne Université, CNRS-ENS-PSL Université, Collège de France, 75252 Paris, France
14Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, United Kingdom
15Faculty of Aerospace Engineering, Delft University of Technology, Kluiverweg 1, 2629 HS Delft, Netherlands

(Received 27 January 2022; revised 10 March 2022; accepted 30 March 2022; published 14 September 2022)

The MIRCOSCOPE mission was designed to test the weak equivalence principle (WEP), stating the equality between the inertial and the gravitational masses, with a precision of $10^{-15}$ in terms of the Eötvös ratio $\eta$. Its experimental test consisted of comparing the accelerations undergone by two collocated test masses of different compositions as they orbited the Earth, by measuring the electrostatic forces required to keep them in equilibrium. This was done with ultrasensitive differential electrostatic accelerometers onboard a drag-free satellite. The mission lasted two and a half years, cumulating five months worth of science free-fall data, two-thirds with a pair of test masses of different compositions. First observed by Galileo and Newton and testing the weak equivalence principle (WEP), according to which all bodies fall in the same way in a gravitational field when no other forces are acting on them, independently of their masses and internal constitutions. First observed by Galileo and Newton and tested by Eötvös et al. at the $5 \times 10^{-9}$ level [20], the

General relativity (GR) offers a remarkable description of gravitational interactions, successfully tested in the anomalous precession of the perihelion of Mercury, the bending of light in a gravitational field, the gravitational redshift, the Shapiro time delay and the change in the periods of binary pulsars from the emission of gravitational waves [1–10]. Gravitational waves from the coalescence of neutron stars and very massive black holes have been observed recently, providing evidence for the existence of black holes and ruling out many beyond-GR models [11–19].

A building block of general relativity is the equivalence principle (EP), according to which all bodies fall in the same way in a gravitational field when no other forces are acting on them, independently of their masses and internal constitutions. First observed by Galileo and Newton and tested by Eötvös et al. at the $5 \times 10^{-9}$ level [20], the
universality of free fall was elevated to a principle by Einstein, the weak equivalence principle (WEP), taken as a cornerstone of general relativity [21].

Still the above tests of GR are classical, i.e., not involving quantum physics. But one does not know how to cast GR into a consistent quantum theory, even if several approaches, including most notably string theories [22], have been developed to tackle this problem. At the conceptual level, also, physicists have been dreaming of a unified theory including strong, electromagnetic, and weak interactions as well as gravity.

As GR cannot be considered as a complete theory, and in view of other questions such as the nature of dark energy [23,24] and dark matter [25,26], it is important to test as precisely as possible the EP. Both the need to complete GR if it is to be turned into a satisfactory quantum theory, and the desire for a unified description of interactions, lead to consider the possibility of new long-ranged interactions and forces, that could lead to very small apparent violations of the EP.

Thereby, a test of the EP appears as a test of one of the basic principles of GR, and also serves as a search for new interactions. While gravity is supposed to be mediated by the (still hypothetical) spin-2 graviton, extremely weak new forces could be mediated by very light or massless spin-0 or spin-1 bosons, that may be thought of as part of a completion of GR. Such EP-violating spin-0 bosons, like the dilaton or other dilaton-like particles, tend to appear within string theories [27–37]. A spin-1 boson U associated with an extension of the standard model gauge group is expected to couple to a combination of baryonic, leptonic (or B-L within grand unification) and electromagnetic currents (with possibly axial couplings, of no effects here) [38,39].

The WEP has been intensively tested throughout the past four centuries [1,20,40–43], and verified to a precision of $2 \times 10^{-13}$ in the first decade of the 21st century [44–46]. This precision has been increased by order of magnitude ($2 \times 10^{-14}$) with the first results of the MICROSCOPE mission in 2017 [47,48], taking advantage of space quietness [49] and of new instrument capabilities [50]. At about the same time, over 48 years of lunar laser ranging (LLR) data allowed for a $7.1 \times 10^{-14}$ precision [51]. We report here the final MICROSCOPE mission results, setting the tightest bound on the validity of the WEP achieved to date, also providing improved constraints on additional new forces [52–54].

The MICROSCOPE space mission was designed following developments of the satellite test of the equivalence principle (STEP) [55,56], built from Chapman’s seminal proposal to test the WEP in the Earth orbit as early as the 1970s [49]. Albeit less ambitious than STEP, it was designed to test the WEP in space in terms of the Eötvös ratio

$$\eta_{A,B} = \frac{a_A - a_B}{a_A + a_B}$$

(1)

$$\approx \left( \frac{m_g}{m_i} \right)_A - \left( \frac{m_g}{m_i} \right)_B = \delta(A,B),$$

(2)

where $a_A$ and $a_B$ are the accelerations of two free-falling test bodies $A$ and $B$, and $m_g$ and $m_i$ their gravitational and inertial masses, respectively. Equation (2) defines the approximated Eötvös ratio to be estimated by MICROSCOPE [57,58].

The bodies are two concentric hollow cylindrical test masses controlled with electrostatic forces in a differential accelerometer. Any difference in the forces required to keep the two test masses in relative equilibrium would provide evidence for an apparent violation of the WEP, originating from an intrinsic violation, or as an effect of extremely small new forces [59]. MICROSCOPE includes two such differential accelerometers called sensor units: in the first one (SUREF), the two test masses have the same composition (PtRh alloy); in the second one (SUREF), the two test masses have different compositions [PtRh(90/10) and TiAlV(90/6/4) alloys]. The former serves as a reference instrument, while the latter is used to test the WEP. The test masses’ characteristics are summarized in Table I, and details about the instrument can be found in Refs. [47,48,60].

The payload was integrated in a drag-free CNES microsatellite able to provide the experiment with a very quiet environment [61]. MICROSCOPE was launched from Kourou on April 25, 2016 and set into a sun-synchronous, dawn-dusk orbit to optimize its thermal stability. The mission ended on October 18, 2018. Reference [62] presents the mission scenario.

The experimental observable relevant to the test of the WEP is the difference between the electrostatic accelerations exerted on the inner (labeled 1) and the outer (labeled 2) test masses of a given sensor unit $\Gamma_d^{\text{meas}} \equiv \Gamma_1^{\text{meas}} - \Gamma_2^{\text{meas}}$. It is directly related to the Eötvös ratio $\eta(2,1) \approx \delta(2,1)$ and to the

### Table I. Main test-mass physical properties measured in the laboratory before integration in the instrument.

<table>
<thead>
<tr>
<th>Measured parameters at 20°C</th>
<th>SUREF</th>
<th>SUREF</th>
<th>SUEP</th>
<th>SUEP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inner mass</td>
<td>Outer mass</td>
<td>Inner mass</td>
<td>Outer mass</td>
</tr>
<tr>
<td></td>
<td>Pt/Rh</td>
<td>Pt/Rh</td>
<td>Pt/Rh</td>
<td>Ti/Al/V</td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>0.401 533</td>
<td>1.359 813</td>
<td>0.401 706</td>
<td>0.300 939</td>
</tr>
<tr>
<td>Density (g cm$^{-3}$)</td>
<td>19.967</td>
<td>19.980</td>
<td>19.972</td>
<td>4.420</td>
</tr>
</tbody>
</table>
various forces acting on the satellite (see Ref. [59] for a
detailed derivation). In the instrument’s reference frame,
\[ \Gamma_{d} \approx \mathbf{K}_{0,d} + \Gamma_{c}^{\text{app}} + 2\mathbf{A}_{d}^{\text{l}} \mathbf{\Gamma}_{\text{ad}}^{\text{app}} + 2\mathbf{A}_{d}^{\text{r}} \mathbf{\Gamma}_{\text{ad}}^{\text{app}} + \mathbf{\Gamma}_{\text{ad}}^{\text{app}} + 2\mathbf{A}_{d}^{\text{r}} \mathbf{\Gamma}_{\text{ad}}^{\text{app}} + \mathbf{\Gamma}_{\text{ad}}^{\text{app}}. \]

where \( \mathbf{K}_{0,d} \) is a differential bias, \( \mathbf{\Delta} \) connects the center of the
inner mass to that of the outer mass, \( \mathbf{T} \) is the gravity gradient
tensor in the satellite frame, \( \mathbf{\Gamma}_{\text{c}} \) is the gradient of inertia, with \( \mathbf{\Omega} \) the satellite’s angular velocity
matrix, \( \mathbf{\Gamma}_{\text{ad}} \) is the Earth’s gravity acceleration at the center of
the satellite, \( \mathbf{\Gamma}_{\text{ad}}^{\text{app}} \) is the mean acceleration applied on both
masses, \( \mathbf{A}_{d}^{\text{l}} \) and \( \mathbf{A}_{d}^{\text{r}} \) are defined from the instrument’s scale factors and test-mass
reference frame defects. Those parameters are more fully
described in Table II of Ref. [47] and in Refs. [48,59,63].

The test of the WEP is performed along the longitudinal
axis of the test masses, designed to be the most sensitive.
Data analysis thus deals with the differential acceleration
(3) projected along the longitudinal, sensitive x axis of the
instrument. Reference [59] provides the measurement
equation projected on this axis.

The satellite can be spun around the normal y axis to the
orbital plane and oppositely to the orbital motion in order to
increase the frequency of the Earth’s gravity modulation. In
this case, in the satellite frame, the Earth’s gravity field
rotates at the sum of the orbital and spin frequencies. A WEP violation would give a signal modulated at this
frequency, denoted \( f_{\text{EP}} \). The frequencies used during the
mission are listed in Table II. Reference [62] introduces the
concept of sessions: WEP test sessions last several days and
are defined by their spin frequencies, while calibration
sessions are short and allow us to estimate instrumental
parameters.

The analysis of first results in Refs. [47,48] used one
120-orbit session on SUEP to obtain \( \delta(T_{i}, P_{i}) = [-1 \pm
9(\text{stat}) \pm 9(\text{syst})] \times 10^{-15} \) at 1σ statistical uncertainty. No
calibration was used, and systematic errors were dominated
by thermal effects, for which an upper bound was used.

**TABLE II.** Frequencies of interest.

<table>
<thead>
<tr>
<th>Frequency ((x \times 10^{-3} \text{ Hz}))</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{\text{orb}} ) ( = 0.168 )</td>
<td>Mean orbital frequency</td>
</tr>
<tr>
<td>( f_{\text{spin}} ) ( = 0.756 \times 10^{-3} )</td>
<td>Spin rate frequency 2 (V2 mode)</td>
</tr>
<tr>
<td>( f_{\text{spin}_3} ) ( = 2.943 \times 10^{-3} )</td>
<td>Spin rate frequency 3 (V3 mode)</td>
</tr>
<tr>
<td>( f_{\text{EP}_2} ) ( = 0.924 \times 10^{-3} )</td>
<td>EP frequency in V2 mode</td>
</tr>
<tr>
<td>( f_{\text{EP}_3} ) ( = 3.111 \times 10^{-3} )</td>
<td>EP frequency in V3 mode</td>
</tr>
<tr>
<td>( f_{\text{cal}} ) ( = 1.228 \times 10^{-3} )</td>
<td>Calibration frequency</td>
</tr>
</tbody>
</table>

Improvements were then expected in the pursuit of the
mission and its data analysis.

We report here the final results of the MICROSCOPE
mission, based on eighteen sessions for SUEP and nine
sessions for SUREF, with all data calibrated and systematics
now fully characterized [63,64]. Reference [65] presents the
methods used for data analysis, and Ref. [66] details their
results on the actual data. The main aspects of the data and its analysis are summarized below.

A handful of sessions were discarded because of non-
linearities at the beginning of the mission, before the
control loop’s electronics was upgraded. A few others
were discarded because of rare anomalies. The results
presented in this Letter were then obtained from eighteen
sessions on SUEP and nine on SUREF, including two with
both SUs switched on together [66]. Hereafter, we first
describe the data analysis methods before discussing the
actual in-flight estimation of instrumental parameters and
presenting the Eötvös parameter determination.

Each in-flight calibration session is dedicated to estimating
one or two parameters and designed so that the signals
sourced by those parameters have a favourable signal-to-
noise ratio. Using a Markov chain Monte Carlo (MCMC)
technique on nine sessions, we showed that it is possible to
cumulate sessions and estimate all parameters simultaneously from Eq. (3). This unpublished study, based on
Ref. [67], used the data of Ref. [47] and gave results
consistent with those of the technique described below. Indeed, instead of using a CPU-expensive MCMC method,
we use the fact that parameters are almost independent to
simplify and better control the estimation process. This is
done via the following iterative method based on the ADAM
(Accelerometric Data Analysis for MICROSCOPE) code
to estimate parameters in the frequency domain. The
method is presented and applied to numerical simulations
in Ref. [65].

When projected on the x axis, the measurement equation
(3) is of the form \( \Gamma_{d,x} = f(p_{k}, t) + n_{d,x} \), where \( p_{k} \)
are parameters and the time dependence is related to measured
or modeled signals \( s(t) \). For each session, the data provide
us with \( \Gamma_{d,x} \) and \( s(t) \), allowing for the estimation of the
parameters \( p_{k} \). Moreover, a \textit{a priori} values \( p_{k,0} \) (either
measured on ground or estimated during an earlier in-flight
calibration session) are used to correct the measurement
for the corresponding signals and to refine the
estimation of some parameters \( p_{k} \).

In practice, instrumental defects are parametrized by the
\( \mathbf{K}_{0,d} \) and \( \mathbf{\Delta} \) vectors, as well as the \( \mathbf{A}_{d}^{\text{l}}, \mathbf{A}_{d}^{\text{r}}, \) and \( \mathbf{C}_{d} \)
matrixes in Eq. (3). Note that only some of their components
impact the projected acceleration. Reference [59] details how they affect the measurement and Ref. [63]
describes how they were estimated in flight and evolve in time.

The estimation of \( \Delta_{x} \) and \( \Delta_{z} \) takes advantage of their
couplings with the Earth’s gravity gradient, whose strong
line at \( 2f_{\text{EP}} \) allows for a direct determination in science data.
from an accurate Earth’s gravity model [68,69]. Dedicated 5-orbit sessions were used to measure $\Delta_y$, where the satellite was oscillated about the z axis at frequency $f_{\text{cal}}$ to create a measurable signal driven by $\Delta_y$ at $f_{\text{cal}}$. The elements of the first row of the $[A_d]$ matrix $a_{dji}$ were measured by shaking the satellite at frequency $f_{\text{cal}}$ along each axis (x to measure $a_{d11}$, y for $a_{d12}$, and z for $a_{d13}$) in order to drive a measurable signal dependent on those parameters. The $a_{d11}$ sessions also allowed for a measurement of the differential quadratic factor $K_{2d,xx}$ at $2f_{\text{cal}}$. Although we found a slight correlation of $\Delta_y$ and $\Delta_z$ with temperature, the other parameters remained roughly constant during the mission [63]. Table III lists their mean values.

Once the above iterative process had converged (typically in two to three iterations), we estimated the Eötvös parameter on calibrated data following the corrected measurement equation

$$\Gamma_{d,xx}^{\text{corr}} = \tilde{b}_{d,xx} + \delta_x g_t + \delta_z g_t + \Delta_y S_{xx} + \Delta_z S_{xz} + n_{d,xx},$$

which is the core model fitted to the data, where $\tilde{b}_{d,xx}$ is the bias. In addition to the Eötvös parameter $\delta_x$ we also estimated the amplitude $\delta_z$ of a signal proportional to $g_z$ (varying also at $f_{\text{EP}}$ but in quadrature with $g_x$) and the components $\Delta_y$ and $\Delta_z$ of the apparent off centering.

Additionally, spurious events can be spotted in the data. “Glitches” are short-lived events, most probably due to crackles of the satellite’s multilayer insulator [70]. Glitches occur quasiperiodically and can impart a signal at $f_{\text{EP}}$. Although numerical models do not allow for the estimation of the level of this signal, we noticed that removing glitches from sessions with a strong signal at $f_{\text{EP}}$—statistically inconsistent with other sessions—decreases the signal, hinting at a significant effect from glitches on the Eötvös parameter estimation. Therefore, to counteract their direct effect on MICROSCOPE’s WEP measurement, we masked them as follows. We used a standard recursive $\sigma$-clipping technique—$\sigma$ being the standard deviations of the data—to search for outliers, defined as points that deviate by more than (i) $4.5\sigma$ from the moving average of the data and (ii) more than $3\sigma$ from the moving average of the high-frequency-filtered data. We then mask one (15) second(s) before (after) each outlier to make sure that the transient regime was always removed [62,66]. Masked glitches thus behaved as “missing” data, so that data became unevenly sampled in time, thereby hampering ADAM’s fit in the frequency domain [71,72].

We tackled this difficulty with the M-ECM (modified-expectation-conditional-maximization [73]) technique; it maximizes the likelihood of available data through the estimation of missing data by their conditional expectation, based on the circulant approximation of the complete data covariance. We showed in Ref. [73] that it faithfully reconstructs the noise power spectral density and provides unbiased estimates of parameters. Finally, we added and correctly measured mock WEP violation signals in the data to make sure that this procedure does not affect a possible real WEP violation signal.

M-ECM also fills gaps, and we can then use ADAM to cross-check M-ECM’s estimates of the Eötvös parameter for each session. We also use it to combine all sessions and infer the overall constraint given below.

In addition to glitches, rare jumps in the differential acceleration can be spotted, mostly on SUREF [66]. These jumps are not simple discontinuities, but appear as unsteady transitions between two stable states. Although hidden in the noise, they perturb the data analysis and must be discarded. Since this amounts to creating gaps of several hundred seconds, we decided to extract “segments” between jumps (or between jumps and any extremity of the session), when such jumps existed (otherwise, we call segment the entire session). Segments are as long as possible and consist of an even number of orbital periods to ensure that potential contamination by signals at frequencies $mf_{\text{orb}} + nf_{\text{spin}}$ ($m, n \in \mathbb{N}$) are canceled [66].

Figure 1 shows the estimates of the Eötvös parameter for each segment, obtained with M-ECM (blue circles) and ADAM (orange diamonds). The two methods are perfectly consistent. Error bars vary in accordance with the duration of segments and with the spin rate: the higher the spin rate, the lower the error bars, since the noise is minimal for the highest spin rate, see Ref. [66]. The black lines and gray areas show the combined constraints and their 68% confidence region [65,66]. $\delta(\text{Pt, Pt}) = (0.0 \pm 1.1) \times 10^{-15}$ for SUREF and $\delta(\text{Ti, Pt}) = (-1.5 \pm 2.3) \times 10^{-15}$ for SUEP. Those uncertainties contain statistical errors only. We discuss systematic errors below.

We found that the overall systematics upper bound is $2.3 \times 10^{-15}$ for SUREF and $1.5 \times 10^{-15}$ for SUEP, compared to specifications of $0.2 \times 10^{-15}$ [63]. Except non-linearity (as discussed in Ref. [64]) and temperature variations (see below), all contributors to the systematics error budget have effects lower than required. For instance, the contribution of the Earth gravity gradient could be canceled by the precise estimation of the test masses off

---

**TABLE III. Mean estimated values of the off-centerings’ components and of the first row of the $[A_d]$ matrix estimated in flight [63].** The off-centerings’ components $\Delta_x$ and $\Delta_z$ correspond to sessions within a limited temperature range.

<table>
<thead>
<tr>
<th></th>
<th>SUEP</th>
<th>SUREF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_x$ ($\mu m$)</td>
<td>$19.998 \pm 0.009$</td>
<td>$-35.884 \pm 0.005$</td>
</tr>
<tr>
<td>$\Delta_y$ ($\mu m$)</td>
<td>$-8.19 \pm 0.09$</td>
<td>$5.89 \pm 0.05$</td>
</tr>
<tr>
<td>$\Delta_z$ ($\mu m$)</td>
<td>$-5.605 \pm 0.009$</td>
<td>$5.712 \pm 0.005$</td>
</tr>
<tr>
<td>$a_{d11}$ ($10^{-3}$)</td>
<td>$8.5 \pm 0.2$</td>
<td>$-14.6 \pm 0.2$</td>
</tr>
<tr>
<td>$a_{d12}$ ($10^{-5}$ mrad)</td>
<td>$-25.6 \pm 0.5$</td>
<td>$-3.5 \pm 1.5$</td>
</tr>
<tr>
<td>$a_{d13}$ ($10^{-5}$ mrad)</td>
<td>$13.6 \pm 0.9$</td>
<td>$-9.1 \pm 0.3$</td>
</tr>
<tr>
<td>$K_{2d,xx}$</td>
<td>$-1037 \pm 4800$</td>
<td>$2409 \pm 1650$</td>
</tr>
</tbody>
</table>
centerings. Local gravity effects were mitigated by a careful design of the satellite and of the instrument [63]. Similarly, we showed in Ref. [63] that magnetic effects, due to the interaction of the test masses with the Earth’s magnetic field, were well below the requirements, as expected from the integration of a magnetic shield around the payload. The DFACS performance was much better than expected, allowing for a residual linear accelerations at $f_{\text{EP}}$ smaller than $2 \times 10^{-13} \text{ ms}^{-2}$ ($6 \times 10^{-14} \text{ ms}^{-2}$) in (out of) the orbital plane, and for residual angular accelerations smaller than $2.5 \times 10^{-11} \text{ rad s}^{-2}$; the related systematic errors are well below the specifications.

Temperature variations are the main source of systematic errors. They induce a differential acceleration through a thermal sensitivity of the SU and of the front end electronics unit (FEEU). Specific sessions were designed to characterize the thermal sensitivity through a periodic stimulus by on-board heaters; the temperature and differential accelerations were finely monitored and compared to provide better estimates of the thermal sensitivities at different frequencies [63,74]. We found a linear frequency dependence for SUEPs thermal sensitivities, but none for SUREFs.

On the other hand, the temperature data during EP sessions only allowed for a pessimistic upper bound since the temperature variations were smaller than the temperature probes noise at $f_{\text{EP}}$. In response to this limitation, additional sessions were devoted to confirm the thermal design of the satellite, in order to show that temperature variations are driven by the Earth’s albedo coming through the FEEU radiator’s baffle (Fig. 15 of Ref. [63]), inducing a modulation of the temperature at $f_{\text{EP}}$ [63]. A first session, based on heating the FEEU panel with local heaters, allowed us to show that the impact of the Earth’s albedo on the satellite walls is negligible. In a second session, the satellite was tilted by 30° about its spin axis in inertial mode during 465 orbits ($32.3$ days) in order to maximize the albedo light entering the FEEU radiator. We found that temperature variations are attenuated by a factor 500 between the FEEU and the SU. Based on the data available (at frequencies lower than $10^{-3} \text{ Hz}$), we took this factor 500 as the lowest limit to compute an upper bound of temperature fluctuations at frequencies higher than $10^{-3} \text{ Hz}$, in particular at $f_{\text{EP}}$, where temperature probes allow for a measurement of the FEEU temperature variations but not of the SU’s, since it is below the probe’s noise [63].

Putting these results together, $\text{MICROSCOPE}$’s new constraint on the validity of the WEP is

$$\eta(\text{Ti}, \text{Pt}) = \left[ -1.5 \pm 2.3(\text{stat}) \pm 1.5(\text{syst}) \right] \times 10^{-15},$$

where the statistical error is given at 1σ, and where we identified the measured, approximated Eötvös ratio $\delta$ with the exact one $\eta$. This result is close to the $10^{-15}$ precision for which the mission was designed, and improves our previous constraints [47] by a factor 4.6. The reference instrument provided a null result, $\eta(\text{Pt}, \text{Pt}) = [0.0 \pm 1.1(\text{stat}) \pm 2.3(\text{syst})] \times 10^{-15}$, showing no sign of unaccounted systematic errors in Eq. (5). As expected, SUREF’s statistical error is smaller than SUEP’s because it is more sensitive.

Beside constraining the validity of the WEP to an unprecedented level, $\text{MICROSCOPE}$ also allows for unprecedented constraints on topics as various as Lorentz invariance [75], long-range interactions [52–54,76,77], or dark matter searches [78]. It also paves the way to new, more ambitious experiments to test GR in space [79]. The analyses presented in this Letter and in Refs. [47,48] provide essential feedback for future upgrades on the payload and satellite sides that can lead to the next-generation $\text{MICROSCOPE}$ mission. In particular, the gold wire allowing for the test masses charge management should be replaced by a
contactless device, such as the one proven to work in space in LISA Pathfinder [80–82]. Glitches should be reduced to a minimum, e.g., by tightening the requirement on the crackles of the satellite’s coating, or their effect should be better understood, e.g., through a better understanding of the full transfer function of the satellite instrument system, so that it can be efficiently corrected for. Furthermore, a better thermal stability and thermal characterization of the system will allow us to beat the thermal systematics. With these upgrades, not only should it be possible to reach a $10^{-15}$ precision on the Eötvös ratio, but also to provide unprecedented constraints on topics as various as Lorentz invariance [77], long-range interactions [52,53], or dark matter searches [78].

The authors express their gratitude to all the different services involved in the mission partners and in particular CNES, the French space agency in charge of the satellite. This work is based on observations made with the T-SAGE instrument, installed on the CNES-ESA-ONERA-CNRS-ONERA MICROSCOPE mission. ONERA authors work is financially supported by CNES and ONERA fundings. Authors from OCA, Observatoire de la Côte d’Azur, have been supported by OCA, CNRS, the French National Center for Scientific Research, and CNES. ZARM authors work is supported by the DLR, German Space Agency, with funds of the BMWi (FKZ 50 OY 1305), and by the Deutsche Forschungsgemeinschaft DFG (LA 905/12-1). The authors would like to thank the Physikalisch-Technische Bundesanstalt institute in Braunschweig, Germany, for their contribution to the development of the test masses with funds of CNES and DLR.

*Deceased.
††Gilles.Metris@oca.eu
manuel.rodrigues@onera.fr
¶Present address: CEA, Centre de Saclay, IRFU/DPhP, 91191 Gif-sur-Yvette, France.
**Present address: DLR, Institute for Satellite Geodesy and Inertial Sensing, Am Fallturm 9, D-28359 Bremen, Germany.
†††Present address: AIRBUS Defence and Space, F-31402 Toulouse, France.
‡‡‡Present address: KINEIS, F-31520 Ramonville Saint-Agne, France.
§Present address: GERADTS GMBH, Kleiner Ort 8, D-28357 Bremen, Germany.


[66] G. Métris (Microscope team), MICROSCOPE WEP test and result, Classical Quantum Gravity (to be published).


[80] M. Armano et al., Constraints on LISA Pathfinder’s self-gravity: Design requirements, estimates and testing procedures, Classical Quantum Gravity 33, 235015 (2016).
