# MICROSCOPE: past, present and future

Gilles Metris, Manuel Rodrigues, Joel Bergé

# MICROSCOPE: past, present and future

G. MÉTRIS
*Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, IRD, Géoazur, F-06560 Valbonne, France*

M. RODRIGUES, J. BERGÉ
*DPHY, ONERA, Université Paris Saclay, F-92322 Châtillon, France*

The MICROSCOPE mission tested the Weak Equivalence Principle (WEP) with an unprecedented precision of order $10^{-15}$, two orders of magnitude better than the previous best lab experiments. While the WEP, the cornerstone of General Relativity (GR), does not sway, the decade-long problems faced by fundamental physics stay still: how can we unify GR with the Standard Model, and how can we explain the acceleration of the cosmological expansion? As most beyond-GR models predict a violation of the WEP, albeit at an unknown level, it remains critical to even better test the WEP. In this paper, we review the MICROSCOPE mission, give its final constraint on the WEP, and build on its experimental limitations to show how we could improve them by a further two-order of magnitude in the precision of the test of the WEP.

## 1 Introduction

The universality of free-fall (UFF) has been recognized since Galileo rolled objects down inclined planes and found that, locally, they all undergo the same gravitational acceleration: all objects within the same gravitational field fall at the same rate, independently of their mass and composition. With Newton's second law, the UFF can be restated as the proportionality between the gravitational mass $m_G$ and the inertial mass $m_I$, with the same proportionality constant for all bodies: this is the usual definition of the weak equivalence principle (WEP).

In the early 20th century, Einstein generalized the WEP, stating that in small enough regions of spacetime, the non-gravitational laws of physics reduce to those of special relativity. In particular, one cannot detect a gravitational field by means of local experiments. This is known as the Einstein equivalence principle (EEP). A subsequent version, the strong equivalence principle, generalizes the EEP to gravitation. That was the starting point to general relativity (GR).

GR describes gravitation as the simple of spacetime's curvature, while recovering Newton's description of gravitation as a classical inverse-square law force in weak gravitational fields and for velocities small compared to the speed of light. As a highly predictive theory, it has so far successfully passed all experimental tests [1]. Standing next to GR, the Standard Model (SM) was built from the realization that the microscopic world is intrinsically quantum.

Although both GR and SM leave few doubts about their validity in their respective regimes, scientists have been faced with difficulties for decades. Firstly, the question of whether GR and the SM should and could be unified remains open: major theoretical endeavors delivered models such as string theory, but still fail to provide a coherent vision of the world. Secondly, the unexpected dark matter and dark energy make up most of the Universe's mass-energy budget.

The WEP has been tested for four centuries with increased precision [2]. The concept of a

test in space emerged in the 1970s [3], motivated by the quiet environment that space can provide and by the benefit of test periods much longer than on-ground experiments. In 1999, ONERA (Office National d'Etudes et de Recherches Aérospatiales) and OCA (Observatoire de la Côte d'Azur) proposed the MICROSCOPE mission (MICRO-Satellite à Compensation de traînée pour l'Observation du Principe d'Equivalence) to CNES. Selected within the framework of the MYRIADE micro-satellite line, MICROSCOPE operating at room temperature aimed to test the WEP with a more modest accuracy than the space cryogenic missions STEP, QuickSTEP, MiniSTEP or GEOSTEP [4], all of which have been canceled by now.

It was the start of a long, winding path. After a few years of budget freeze, the project entered the core of its development in 2006 and was faced with technical difficulties: change to field-emission electric indium propulsion in 2006 and finally to cold gas propulsion in 2009, breakage of the 7 $\mu$m gold wire used to manage the charge of the payload's test-masses during the qualification in 2011, 1 ns timing anomaly for the payload digital signal processor (DSP) flight model in 2012, coupling in the electrostatic actuation of the payload during satellite integration test in 2015. All these pitfalls were successfully solved by the core CNES/ONERA/OCA team. At the limit of test-ability on ground and at the limit of performance for each subsystem, these difficulties were never encountered before in ONERA's long experience in accelerometry [5], and show how difficult it is to push the limits of the state of the art in a space experiment.

MICROSCOPE was finally launched in 2016. After successfully dealing with unexpected anomalies [6], the mission provided two and a half years of useful data. In 2017, a first analysis based on only 7% of the eventual science data allowed us to verify the WEP at $2 \times 10^{-14}$ sensitivity level [7,8]. In 2022, the full data allowed us to improve that precision by one order of magnitude [9,10]. In this paper, we first describe the MICROSCOPE in Sect. 2. In Sect. 3, we describe the data processing and provide the upper bound provided by MICROSCOPE on the validity of the WEP. Finally, Sect. 4 draws ways of improvements to design a future advanced MICROSCOPE mission.

## 2 MICROSCOPE mission overview

### *2.1 WEP test experiment principle*

The measurement relies on comparing the accelerations of two concentric bodies – cylinders in the case of MICROSCOPE – in orbit around the Earth. As shown in Fig. 1, the measurement is performed along the cylinders' X-axis, which is aligned with their main axis. In an inertial pointing configuration, it is pointing in the same direction of the Earth's gravity field vector once per orbit. In a perfect case, the difference of acceleration is proportional to the Eötvös parameter defined by the relative ratio of difference of gravitational-to-inertial masses $m_{gj}/m_{ij}$ between two materials $j$:

$$\delta(2,1) = 2\frac{a_2 - a_1}{a_2 + a_1} = 2\frac{m_{g2}/m_{i2} - m_{g1}/m_{i1}}{m_{g2}/m_{i2} + m_{g1}/m_{i1}}, \tag{1}$$

where $a_j$ are the acceleration undergone by the two bodies.

In MICROSCOPE, the test-masses are part of a double concentric accelerometer. The test-masses are finely controlled by electrostatic forces to be motionless with respect to the surrounding electrodes as illustrated in the right panel of Fig. 1. The forces applied by the set of electrodes are determined by the voltage applied on the test-mass and on each electrodes[11]. The combination of these voltages with the geometry of the instrument defines the electrostatic forces and torques applied to each test-mass in order to counteract all the other effects that prevent the test-mass to stay motionless with respect to the satellite.

Thus, if a WEP violation exists, it can be detected as a signal with a well-known frequency (the orbital frequency $f_{\rm orb}$ in the case of Fig. 1) in the differential acceleration measured by the accelerometer (i.e., the difference of electrostatic force per unit mass between the two test

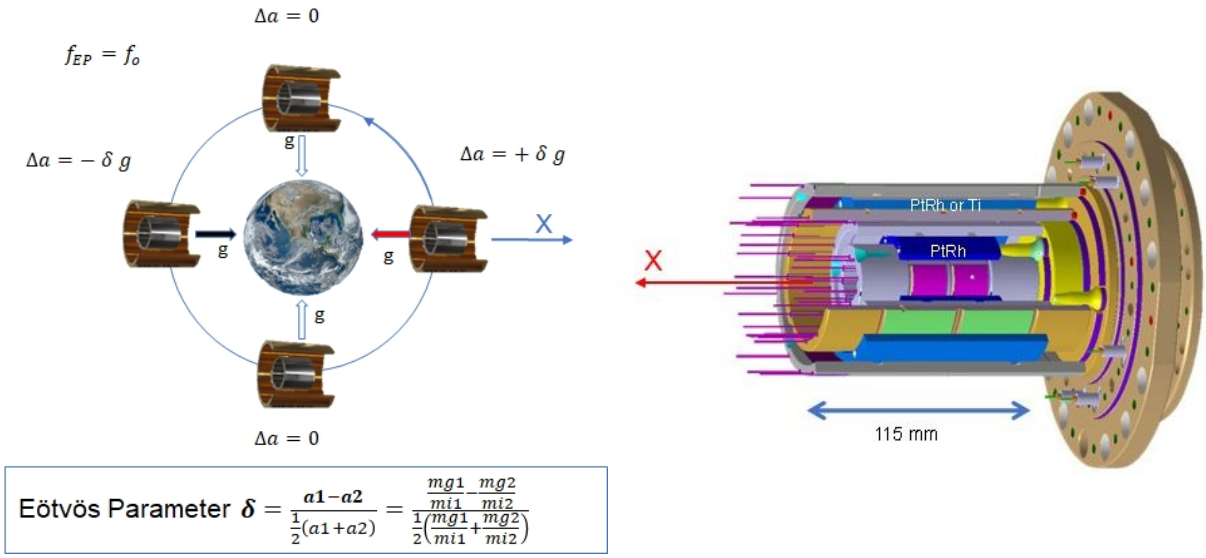


Figure 1 – Experimental principle (left) and accelerometer core (right).

masses). The measurement precision can be improved by rotating the satellite about the axis normal to the orbital plane. This increases the modulation frequency of the Earth's gravity vector projected onto the X-axis, to put it closer to the minimum of the instrumental noise. The WEP-violation frequency becomes $f_{\rm EP} = f_{\rm orb} + f_{\rm spin}$, with $f_{\rm spin}$ the rotation frequency of the satellite. Two spin frequencies have been used during the mission, leading to two test measurement data sets at $f_{\rm EP} \approx 0.9 \times 10^{-3}$ Hz and $f_{\rm EP} \approx 3.1 \times 10^{-3}$ Hz.

### *2.2 Payload*

The payload [11] is composed of two identical differential accelerometers also called sensor units (SUs) except for the test-mass material. Each SU have two concentric hollow cylindrical test-masses surrounded by electrodes engraved on gold-coated silica parts. Each SU is connected to a front-end electronics unit (FEEU) which delivers the voltages to the test-masses and electrodes and transmits the data to the interface control unit (ICU). Each ICU connected to the FEEU contains all the digital electronics and software to operate the test-mass control servo-loops and data conditioning for the satellite and then the ground telemetry. The SU and the FEEU are integrated in a thermal cocoon placed at the core of the satellite which offers a micro-Kelvin stability around the measurement frequencies.

The first SU, called SUREF, comprises two test-masses of the same material: PtRh10 platinum-rhodium alloy containing 90% by mass of Pt ($A = 195.1$, $Z = 78$) and 10% Rh ($A = 102.9$, $Z = 45$). SUREF is dedicated to experiment and accuracy verification (in orbit or on ground within the data processing) as it is supposed to give a null signal at $f_{\rm EP}$. The second SU, called SUEP, comprises two test-masses of different material: the same PtRh10 alloy for the inner test-mass and an aeronautic titanium alloy (TA6V) for the outer test-mass with the atomic composition 90% of titanium ($A = 47.9$, $Z = 22$), 6% of aluminium ($A = 27.0$, $Z = 13$) and 4% of vanadium ($A = 50.9$, $Z = 23$). SUEP is dedicated to the WEP test.

Each test-mass defines a six-degree-of-freedom accelerometer. In order to operate in the most quiet environment and to get the most accurate orientation of the satellite, the accelerometer outputs are used by the satellite's drag-free and attitude control system (DFACS): it applies the necessary commands to the cold gas thrusters (Fig. 2). Atmospheric drag, Sun and Earth radiation forces, magnetic torques and all other disturbing sources are therefore compensated in order to nullify the common mode of one of the SU (i.e., either the mean acceleration of the two concentric test-masses or one of the acceleration output). The accelerometer's output or its

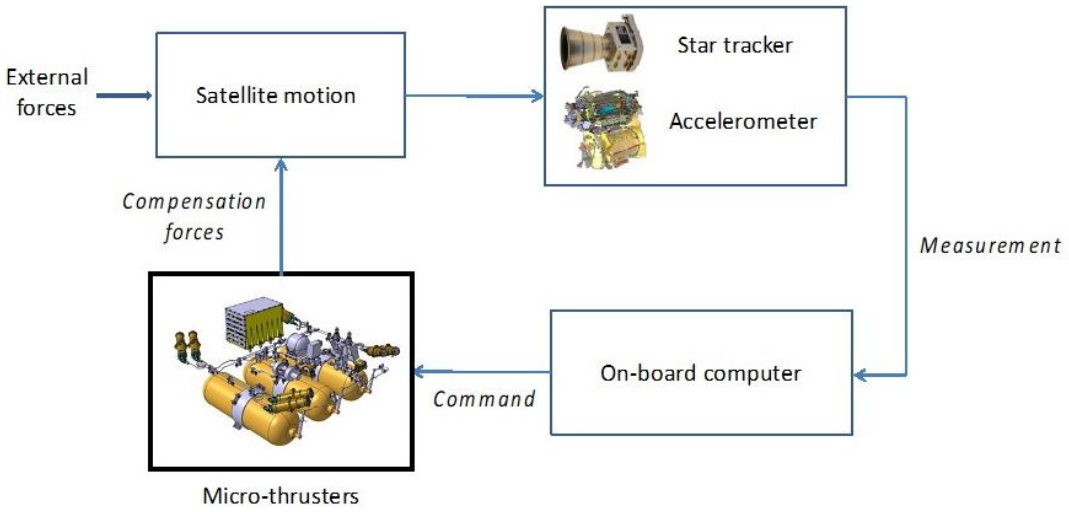


Figure 2 – Satellite Drag-Free and Attitude Control System

internal servo-loop can be artificially biased at a particular frequency to stimulate the test-mass or the satellite (linear or angular motion) during calibration sessions.

### *2.3 Drag-free satellite*

One of the challenges of the mission objectives is to make the satellite environment as quiet as possible for the payload to prevent any corruption of acceleration measurements.

The MICROSCOPE mission has been developed on the basis of scientific missions exploiting the CNES MYRIADE microsatellite product line whose architecture comprises a platform with generic functional chains (energy, communication, computer, structure, etc.). Some adaptations and modifications were necessary to cope with the unusual performance requirements. Usually, the payloads of the MYRIADE satellites are located on the decoupled upper part of the platform but MICROSCOPE payload module has been accommodated at the center of the spacecraft where it could take advantage of a more stable thermal environment (Fig. 3).

The satellite thermal design has been optimised to offer the payload a tight temperature stability: the required stability around the WEP test frequency $f_{\mathrm{EP}}$ was set to 1 mK at the sensor unit interface and to 10 mK at the associated analog electronics interface. Active heaters did not operate during the science operations in order to avoid any interference with the payload measurements. Consequently, the thermal control on the satellite purely relied on passive methods: the dissipation of the electronic units was ensured by satellite external radiators. The in-orbit estimated thermal performance exceeded requirements and expectations. The payload was also shielded from the Earth and satellite magnetic field. In addition, the mechanical or electronic micro disturbances were minimized by a careful design and analysis to ensure an optimal environment: choice of multi-layer insulation (MLI) to minimize cracking, minimisation of current loops, study of thermoelastic deformations to estimate internal gravitational effects...

To counteract non-gravitational forces and torques, an active control of accelerations and attitude of the satellite was implemented through the DFACS (Fig. 2). The DFACS used the scientific instrument itself as main sensor for delivering the linear as well as the angular accelerations hybridized with the star tracker measurements. The control laws for acceleration and attitude estimated the total forces and torques to be applied on the satellite which were transformed into eight micro-thrust commands sent to the cold gas propulsion system placed on two opposite walls of the satellite (Fig. 3). Each of the eight pods of thrusters actually comprises two thrusters: one nominal and one redudant not operating but that could be switched on in case of failure of the nominal. The DFACS in-orbit performances allowed to reduce the disturbances

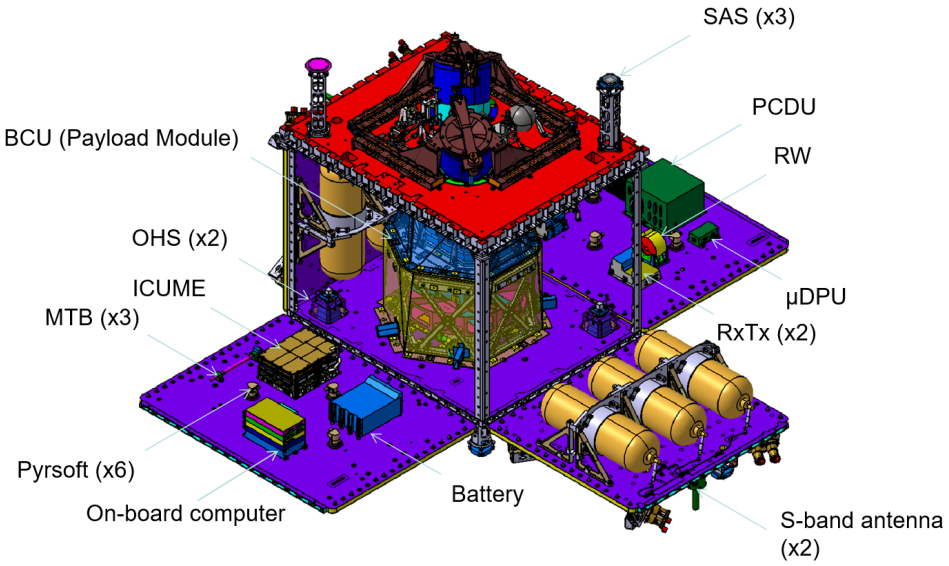


Figure 3 – The cube forming the satellite is open in the picture, the instrument T-SAGE is at the center surrounded by the two $2\times3$ tanks of the cold gas propulsion system. Once closed the satellite cube measures $1.4\,\mathrm{m}\times1\,\mathrm{m}\times1.5\,\mathrm{m}$ and weighs about 300 kg.

by 90 dB around $f_{\mathrm{EP}}$ leading to a controlled linear acceleration better than $3\times10^{-13}\,\mathrm{m\,s^{-2}}$, one order of magnitude better than expectation. The satellite attitude was controlled to better than 1 $\mu$rad at $f_{\mathrm{EP}}$ with an angular velocity stability better than $3\times10^{-10}\,\mathrm{rad\,s^{-1}}$ at $f_{\mathrm{EP}}$ in rotating mode, one order of magnitude better than expectation as well. The induced angular acceleration was controlled to better than $10^{-11}\,\mathrm{rad\,s^{-2}}$ at $f_{\mathrm{EP}}$, limiting centrifugal effects due to the off-centring of the test-masses.

Besides, the DFACS was able to receive additional external sine signals at particular frequencies in order to calibrate the instrument (differential scale factor, test-mass alignments and off-centerings, coupling between axes, non-linearity). Particular sessions were also dedicated to thermal sensitivities (see section 3.3) thanks to dedicated heaters.

## 3 Data processing

### *3.1 Measurement equation*

A single accelerometer (called inertial sensor) measures the difference of acceleration between the test-mass of the accelerometer and the center of mass of the satellite. A differential accelerometer yields the difference $\vec{\Gamma}^{(d)} = \vec{\Gamma}^{(1)} - \vec{\Gamma}^{(2)}$ of two such accelerations for two test-masses. The accelerometers are not perfect, in the sense that we look for very small signals and thus any little defect can make deviate from an ideal response: they have bias, scale factors departing from unity, non-zero coupling between axes [12]. Moreover, their orientation in the satellite, in space and with respect to the Earth's gravity field, is not perfectly known. That is why the measured differential acceleration $\vec{\Gamma}^{(d)}$ is not identical to the real one $\vec{\gamma}^{(d)}$, but is related to it as [12]:

$$\vec{\Gamma}^{(d)} = \vec{b_0}^{(d)} + \left[\mathbf{A}^{(c)}\right]\vec{\gamma}^{(d)} + 2\left[\mathbf{A}^{(d)}\right]\vec{\gamma}^{(c)} + \vec{n}^{(d)}, \tag{2}$$

where

- $\vec{b_0}^{(d)}$ is the difference of bias between the two inertial sensors;
- $\left[\mathbf{A}^{(c)}\right]$ is the common mode sensitivity matrix, close to the identity matrix, which includes scale factors, coupling between axes and global rotation common to the two sensors;

- $\left[\mathbf{A}^{(d)}\right]$ is the differential mode sensitivity matrix, very small, which takes into account the difference of characteristics of the two sensors;
- $\vec{\gamma}^{(c)}$ is the common mode acceleration which is mainly due to non-gravitational accelerations acting on the satellite and not on the enclosed test-masses; these non-gravitational accelerations include drag and radiation pressures and the thrust applied to the satellite which is servo-controlled in order to considerably reduce $\vec{\gamma}^{(c)}$ in the frequency band of interest;
- $\vec{n}^{(d)}$ is the (colored) noise.

In addition, couplings with angular accelerations and nonlinearities can also arise. These terms are not formally included in the above equation but specific measurement sessions have been dedicated to the identification of such effects and demonstrated that they are negligible [13].

The potential WEP-violation signal, $\delta(2,1)\vec{g}$, is included in $\vec{\gamma}^{(d)}$ which also contains the gravity gradient and the differential angular acceleration due to the small residual off-centring between the two test-masses [12]:

$$\vec{\gamma}^{(d)} = \delta(2,1)\vec{g}(O_{\rm sat}) + ([\mathrm{T}] - [\mathrm{In}])\vec{\Delta} + \vec{b_1}^{(d)}, \tag{3}$$

where

- $\vec{g}(O_{\rm sat})$ is the gravity acceleration;
- [T] is the gravity gradient tensor;
- [In] is the gradient of inertia matrix;
- $\vec{\Delta}$ is the off-centring vector from the center of test-mass (1) to the center of test-mass (2);
- $\vec{b_1}^{(d)}$ contains the differences between the other small (mainly non gravitational) perturbations acting on the two test-masses.

Only the axis of the cylindrical test-masses, called $X$, which is much more precise than the other axes is used to estimate the EP signal. Therefore Eq. (2) has to be projected onto the X-axis. This leads to numerous terms [12] but the following considerations lead to simplifications for the reader's convenience:

- the more impacting components of the sensitivity matrix are estimated thanks to dedicated calibrations [14];
- the projection of the common mode is corrected thanks to the calibration of $\left[\mathbf{A}^{(d)}\right]$ and the measurement of $\vec{\gamma}^{(c)}$ (which is roughly assimilated to $\vec{\Gamma}^{(c)}$);
- the effect of the angular acceleration (anti-symmetric part of matrix [In]) is neglected (in practice we can correct for it but we have verified that this has no impact at the $f_{\rm EP}$ frequency thanks to the very good stability of the attitude control);
- small terms as the effect of the out-of-orbital-plane component of the off-centring are corrected thanks to dedicated calibrations;
- the tiny impact of the bias at the $f_{\rm EP}$ frequency is included in the evaluation of the systematic effects.

The remaining model used to analyse the measurements along the X-axis reads

$$\Gamma^{(d)}_{x,\rm corr} = \sum_{j=0}^{3} \alpha_j (t-t_0)^j + \delta_x g_x + \delta_z g_z + \Delta'_x S_{xx} + \Delta'_z S_{xz} + n_x^{(d)}, \tag{4}$$

where

- $\delta_x \approx A^{(c)}_{(1,1)}\delta(2,1)$ ($A^{(c)}_{(1,1)}$ being the scale factor along X) is very close to the Eötvös ratio;
- $\delta_z$, a small fraction of $\delta(2,1)$, is in principle too small to be estimated but is included in the model to check the absence of anomaly;
- $S_{xx}$ and $S_{xz}$ are components of the matrix [S] which is the symmetric part of [T] − [In];
- $\Delta'_x$ (close to $\Delta_x$) and $\Delta'_z$ (close to $\Delta_z$) are "effective" components of the off-centring taking into account the sensitivity matrix;
- $\sum_{j=0}^{3} \alpha_j (t-t_0)^j$ is an empirical polynomial term aiming to absorb the effect of the bias and its slow drift (mainly due to thermal effects).

### *3.2 Data artefacts*

The acceleration measurements were plagued with artefacts. Among them were short instrumental transients ("glitches") rising up to $\sim 10\,\mathrm{nm.s^{-2}}$ and lasting a few seconds in each SU [15]. While the exact origin of the glitch-generating process is unknown, it is correlated with the satellite's position and orientation with respect to Earth, hinting towards a thermal mechanism related to the illumination by the Earth's albedo with some contribution of the Sun, triggering crackles in the MLI coating of the satellite walls.

Due to this correlation, the time distribution of glitches is modulated by the EP frequency. This creates a spurious excess of power in the frequency spectrum, leading to an apparent violation of the EP in some scientific measurement sessions. To prevent this effect from perturbing the test, and in the absence of a proper model accounting for the underlying process, we discarded the data points affected by glitches in the analysis. This masking operation amounts to considering corrupted points as missing data. To avoid any noise frequency leakage related to masking, we use a modified expectation-maximization algorithm (M-ECM), an iterative process which estimates the model parameters together with the missing data [16] until a convergence criterion is reached. The estimation of the Eötvös parameter and the reconstructed periodogram we obtain with M-ECM show that the glitch disturbance is successfully mitigated.

Beside glitches, rare jumps occur in the differential acceleration, mostly on SUREF [10]. They are not simple discontinuities, but appear as chaotic, quickly drifting measurements. Fig. 4 shows three such events, two strong and one weak. Although hidden in the noise, those jumps perturb the data analysis and must be discarded. Since this amounts to creating gaps of several hundred seconds, the use of M-ECM is not justified. Rather, we extract "segments" between jumps (or between jumps and any extremity of the session). In the absence of jumps, we call "segment" the entire session. Segments are made as long as possible and consist of an even number of orbital periods to ensure that potential contamination by signals at frequencies $mf_{\mathrm{orb}} + nf_{\mathrm{spin}}$ ($m, n \in \mathbb{N}$) are canceled [10]: this includes the frequency $f_{\mathrm{EP}}$. Two of these segments are shown in Fig. 4.

### *3.3 Main systematic errors: thermal effect*

In Touboul et al [7] systematic errors were dominated by thermal effects due to a poor knowledge of temperature variations at $f_{\mathrm{EP}}$. Their estimation was performed on 300 contiguous orbits and showed no temperature signal exceeding from the probe noise. Thus this noise was taken as an upper limit giving a 15 $\mu$K temperature variation at $f_{\mathrm{EP}}$ at the SU level and hence a systematic acceleration error of $65 \times 10^{-15}\,\mathrm{m\,s^{-2}}$.

Additional sessions dedicated to temperature sensitivity analysis were eventually performed [14]. In terms of duration, almost 5% of the mission duration was dedicated to the thermal characterisation of the satellite and of the payload compared to the 12% of the time dedicated to the EP test with SUEP and 6% of the time dedicated to SUREF. These particular sessions had

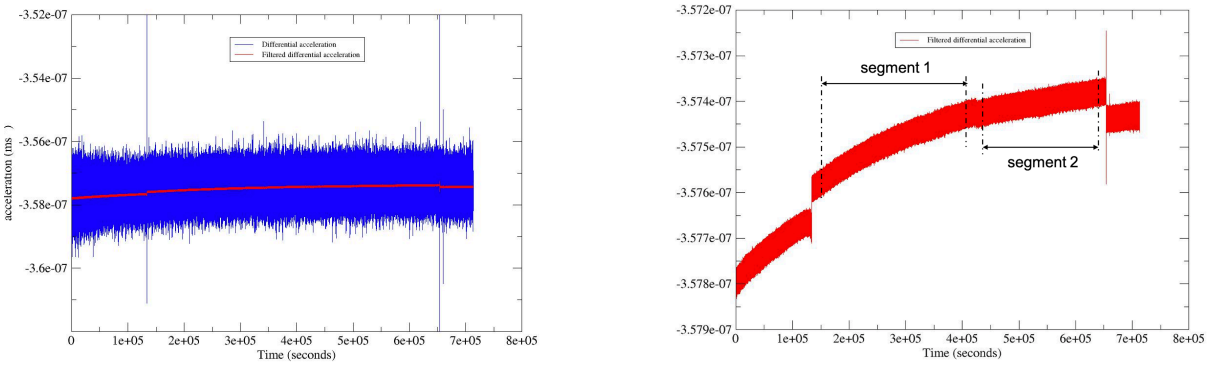


Figure 4 – Example of differential acceleration measured along the $x$-axis. In this case, discontinuities can be seen (most easily in the filtered data, lower panel), and segments are defined according to them. Each segment is analyzed separately. Figure from Ref. [10].

several objectives: (i) evaluate the accelerometer thermal sensitivity model; (ii) confirm that the temperature variations at $f_{\rm EP}$ come from the Earth's albedo entering in the satellite by the FEEU radiator (Fig. 5); and (iii) better evaluate the temperature variations at $f_{\rm EP}$ during the science sessions.

As a first step, to better evaluate the instrument model, the heaters located on the platform at the SU or FEEU level were activated to generate a temperature stimuli and enhance the effect of temperature. Then, calibration sessions were also performed at different temperatures to assess the scale factor dependency on temperature. These experiments led to establish the following model:

$$\begin{aligned}\Gamma^{(d)}_{\rm Tth}(f_{\rm EP}) = &\left[\lambda_{\rm SU}\delta T_{\rm SU}(f_{\rm EP}) + \lambda_{\rm FEEU}\delta T_{\rm FEEU}(f_{\rm EP})\right] \\ &+ \left[\frac{\partial a_{d11}}{\partial T_{\rm SU}}\delta T_{\rm SU}(f_{\rm EP}) + \frac{\partial a_{d11}}{\partial T_{\rm FEEU}}\delta T_{\rm FEEU}(f_{\rm EP})\right]\bar{\Gamma}^{(c)}_{x},\end{aligned} \tag{5}$$

where $\Gamma^{(d)}_{\rm Tth}(f_{\rm EP})$ represents the differential acceleration component of the thermal systematic error at $f_{\rm EP}$, $\lambda_{\rm SU}$ and $\lambda_{\rm FEEU}$ are the differential acceleration sensitivity to the temperature variations of the SUs and FEEU, $a_{d11}$ the scale factor matching and $\Gamma^{(c)}_{x}$ the mean common mode acceleration.

The second step confirmed that the SU's temperature variations were correlated to the FEEU's which follows the temperature of the radiator. The purpose of the FEEU radiator is to evacuate the heat dissipation of the electronics to space. A baffle protects the radiator and limits incoming thermal disturbances from Earth's albedo. More than 460 orbits with a particular inclination of the satellite were performed to amplify the impact of Earth's albedo on the radiator and show that the ratio between FEEU's and SU's temperature variations is higher than 500. Other sessions showed a similar behaviour.

As a conclusion of these thermal tests, it was showed that the disturbance process at $f_{\rm EP}$ comes from the radiator and that the SU's maximal temperature variation in science sessions can be estimated with $\delta T_{\rm SU}(f_{\rm EP}) = \delta T_{\rm FEEU}(f_{\rm EP})/500$. As FEEU's temperature variations come out of the noise in science sessions, it was possible to estimate the SU's maximal temperature variation for each science session at a level lower than 0.1 $\mu$K. By considering this new estimation in Eq. (5), it was possible to estimate the thermal systematics to be lower than $9.3\times10^{-15}\,{\rm m\,s^{-2}}$. This value is now mostly dominated by the FEEU contribution.

### *3.4 Results*

The final results of the MICROSCOPE mission are based on eighteen sessions for SUEP and nine sessions for SUREF [10]. A few sessions were discarded because of non-linearities at the

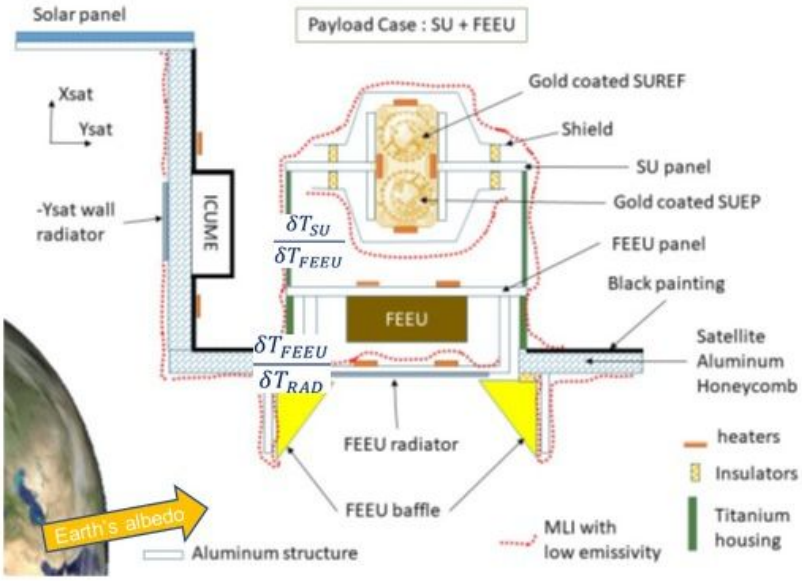


Figure 5 – Payload case and satellite

beginning of the mission (before the control loop's electronics was upgraded) and a few others were discarded because of rare anomalies.

Beside EP-test sessions, in-flight calibration sessions were designed to estimate parameters so that the (perturbing) signals they source have a favourable signal-to-noise ratio (each session being dedicated to one or two parameters). We use the fact that parameters are almost independent to simplify and better control the estimation process with an iterative method based on the ADAM (Accelerometric Data Analysis for MICROSCOPE) code to estimate parameters in the frequency domain [17].

In practice, instrumental defects are parameterized by the $\overrightarrow{b}_1^{(d)}$ and $\overrightarrow{\Delta}$ vectors, as well as the $\left[A^{(d)}\right]$ and $\left[A^{(c)}\right]$ matrices in Eq. (4), with only some of their components impacting the projected acceleration [12,14]. The estimation of $\Delta'_x$ and $\Delta'_z$ uses their couplings with the Earth gravity gradient, whose strong spectral line at $2f_{\rm EP}$ allows for a direct determination in science data based on an accurate Earth gravity model. Dedicated five-orbit sessions were used to measure $\Delta'_y$, where the satellite was oscillated about the $z$-axis at a frequency $f_{\rm cal}$ to create a measurable signal driven by $\Delta'_y$. The elements of the first row of the $[A^{(d)}]$ matrix $a_{d1i}$ were measured by shaking the satellite at frequency $f_{\rm cal}$ 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 $2f_{\rm cal}$. Once the above iterative process converges, the Eötvös parameter is estimated on calibrated data following Eq. (4).

The final MICROSCOPE's constraint on the validity of the WEP is [9]

$$\delta(\mathrm{Ti,Pt}) = [-1.5 \pm 2.3\ (\mathrm{stat}) \pm 1.5\ (\mathrm{syst})] \times 10^{-15}, \tag{6}$$

where the statistical error is given at $1\sigma$.

The reference instrument provided a null result, $\delta(\mathrm{Pt,Pt}) = [0.0 \pm 1.1\ (\mathrm{stat}) \pm 2.3\ (\mathrm{syst})] \times 10^{-15}$, showing no sign of unaccounted systematic errors in Eq. (6). As expected from its higher sensitivity, SUREF's result has a smaller statistical error than SUEP's. On the opposite, it has higher systematic errors (dominated by thermal effects), since they were estimated with less optimal sessions than SUEP's ones [14].

*3.5 Beyond the WEP*

MICROSCOPE was the first space-based laboratory dedicated to testing the WEP [18]. Beside reaching its main goal, the mission also demonstrated technological advances (e.g. 6-degrees-of-freedom drag-free and attitude control) and provided an experience on the limits of the experiment. Furthermore, it allowed for state-of-the-art bounds on ultra-light dark matter and long-range fifth forces [19]. Additionally, it was shown that in principle, a MICROSCOPE-like experiment could constrain short fifth force models and screening mechanisms; nevertheless, MICROSCOPE itself was not designed for those experiments and gave only poor constraints [20].

## 4 The future: MICROSCOPE 2

We can now use MICROSCOPE as a benchmark to plan for improved space tests of the WEP. Currently under preliminary investigation at CNES and ONERA, MICROSCOPE 2 will aim to improve MICROSCOPE's measurement by two orders of magnitude, to reach a $10^{-17}$ precision on the Eötvös parameter. Advanced technological choices, based on the MICROSCOPE's experience and most of them readily available, will allow for this significant jump in precision. In this section, we build on the MICROSCOPE return of experience to envision better instrument and operations, with the goal to gain two orders of magnitude of precision in the test of the WEP in space, while keeping a mission as similar as possible to MICROSCOPE.

In MICROSCOPE, with SUEP and SUREF placed side by side, although stricto-senso a reference instrument, the SUREF instrument could not be used to readily subtract systematic errors from SUEP's measurements. We can cope with this drawback by gathering both instruments. In this way, MICROSCOPE 2 will have only one triple differential accelerometer. The core of the instrument will thus comprise three concentric test-masses, the composition of which must be established according to theoretical considerations. We could either use three different materials (allowing for testing the WEP with more than one pair of materials, as advocated in Fischbach et al [21]), or only two (having two test masses of the same material makes up for an improved SUREF instrument).

MICROSCOPE's stochastic noise was dominated by the thermal dissipation of the gold wires used to control the test masses' charge. Ideally, it should be decreased by at least a factor 100, which could be done by changing the characteristics (length $L$ and diameter $D$) of the wire, since the corresponding noise's spectral density has a $f^{-1/2}$ frequency dependence at low-frequency, the amplitude of which scales [22] as $D^4/L^3$; however, this would imply too long or too narrow a wire, impossible to integrate. On the opposite, a LISA-like charge management system should suppress this limitation [23].

The LISA Pathfinder heritage can also be used to replace MICROSCOPE's capacitive position sensing by a LISA-like interferometer position sensing [24]. This has several advantages: (i) improvement of the acceleration noise at frequency higher than 0.01 Hz, (ii) suppression of the coupling in the loop between the electrostatic action and detection, (iii) possibility to cancel the inertial motion effects (thus relaxing the requirements on the test masses offcentring and on the attitude motion), and (iv) possibility to perform a direct optical differential measurement between two test-masses.

Electrostatic accelerometers, like those of MICROSCOPE, drift in time, which induces low-frequency noise and hampers long measurement sessions. In MICROSCOPE data, it was corrected for with a polynomial fit [17]. Hybridising the electrostatic accelerometer with an atomic interferometer can help improve this, the latter being used as a reference to calibrate the former: despite the increased complexity, cold atoms can bring absolute scale factor determination of the electrostatic accelerometer [25]. This additional instrument is not considered in the baseline but could be eventually added if the technology becomes more mature.

MICROSCOPE data were contaminated by periodic glitches, most likely linked to crackles of the satellite's coating, which can create a signal at the frequency of the WEP test [15]. With

no clear understanding of their physics nor a robust model to quantify their impact, they were discarded from the data. An advanced mission should be designed to prevent them. A solution may be to replace MICROSCOPE's coating with a stiffer structure like that of the GOCE satellite [26].

Operational constraints on the spacecraft limited MICROSCOPE's science sessions to 120 orbits, potentially impacting the experimental conditions, such as temperature, which varies as electronics are turned on and off. An advanced mission should rely on longer sessions to increase the integration time while allowing for more stable experimental conditions. Another operational limitation came from monthly Moon glares on the satellite's star sensor, forcing to depoint the satellite, which induced temperature changes incompatible with the science requirements; additional star sensors that can be switched on and off as needed without acting on the satellite pointing should be considered for MICROSCOPE 2.

Patch fields and contact-potentials differences [27] should be considered in an advanced MICROSCOPE mission (the MICROSCOPE corresponding error budget was evaluated at $2 \times 10^{-13}$ m s$^{-2}$Hz$^{-1/2}$, and was the main limiting factor to constrain a short-range fifth force). To reduce the effect of the contact potential, the gaps between the test-mass and the environment can be increased, since the effect is inversely proportional to the cube of the distance.

Finally, MICROSCOPE showed that increasing the spinning rate of the satellite (to perform the measurement at a higher frequency), improves the temperature stability, as advocated by Nobili and Anselmi [28]. Furthermore, a faster rotation than MICROSCOPE's should allow us to better synchronize the $f_{\rm EP}$ frequency with that of minimal instrumental noise.

## 5 Conclusion

The MICROSCOPE mission has delivered its final measurement on October 2018. Since then, the science team has put a lot of effort into verifying all the data. Some of them were discarded because they were found to be out of specification due to saturation, out-of-performance range (micrometeorite impacts) and non-linearity measurements. Glitches were removed from the data and replaced by a maximum-likelihood noise and signal estimation: this process was verified on simulated violation signal inserted in real data. The estimation of systematic errors have been improved with respect the first results obtained in 2017 to a few $10^{-15}$ in Eötvös parameter units. This allowed for an unprecedented precision on the test of the WEP.

With the lessons learned from the MICROSCOPE mission, we identified key parts of the instrument, satellite and operations that can be improved to beat MICROSCOPE's precision on the test of the WEP by two orders of magnitude. A space mission, as similar as possible to MICROSCOPE (which we tentatively call MICROSCOPE 2) could be designed with only almost off-the-shelf technology. Indeed, most improvements have already been shown to work in space, while others, more demanding, are not compulsory for the success of such as mission. We could thus expect a follow-up mission to MICROSCOPE to fly in the next few decades, on a low enough budget, but with high science outcomes about GR's validity and about ultra-light dark matter.

MICROSCOPE data are available at https://cmsm-ds.onera.fr.

## Acknowledgments

The authors express their gratitude to all the different services involved in the mission partners and in particular the French space agency CNES in charge of the satellite. This work is based on observations made with the T-SAGE instrument, installed on the CNES-ESA-ONERA-CNRS-OCA-DLR-ZARM MICROSCOPE mission. ONERA authors' work is financially supported by CNES and ONERA fundings. Authors from Observatoire de la Côte d'Azur (OCA) have been supported by OCA, the French National Center for Scientific Research (CNRS), and CNES.

ZARM authors' work is supported by the German Space Agency DLR, with funds of the BMWi (FKZ 50 OY 1305 and FKZ 50 LZ 1802) 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 from CNES and DLR.

## References

1. Will, C. M. The Confrontation between General Relativity and Experiment. Living Reviews in Relativity **17**, 4 (2014); Ishak, M. Testing general relativity in cosmology, Living Reviews in Relativity **22**, 1 (2019).
2. Eötvös, L., Pekár, D., Fekete, E., Beiträge zum Gesetz der Proportionalität von Trägheit and Gravität, Ann. Phys. **68**, p. 11 (1922); Bessel, F. W., Versuche über die Kraft mit welcher die Erde Körper von verschiedene Beschaffendheit anzieht, Ann. Phys. Chem. (Poggendorf) **25**, 401-17 (1832); Roll, P. G., Krotkov, R., Dicke, R.H., The equivalence of inertial and passive gravitational mass, Annals Phys. **26**, p. 442-517 (1964); Braginskii, V. B., Panov V. I., Verification of equivalence of inertial and gravitational masses, Zh. Eksp. Teor. Fiz. **61**, 873-879 (1971); Schlamminger, S., Choi, K.-Y., Wagner, T. A., Gundlach, J. H., Adelberger, E. G., Test of the Equivalence Principle Using a Rotating Torsion Balance, *Phys. Rev. Lett.* **100**, 4 (041101)2008; Wagner, T. A., Schlamminger, S., Gundlach, J. H., Adelberger, E. G., Torsion-balance tests of the weak equivalence principle, Class. Quant. Grav. **29**, 18, p. 184002 (2012); Williams, J. G., Turyshev, S. G., Boggs, D. H., Lunar laser ranging tests of the equivalence principle, Classical and Quantum Gravity **29**, 18, p. 184004 (2012); Viswanathan, V., Fienga, A., Minazzoli, O., Bernus, L., Laskar, J., Gastineau, M., The new lunar ephemeris INPOP17a and its application to fundamental physics, MNRAS **476**, p. 1877-1888 (2018).
3. Chapman, P. K. and Hanson, A. J., An Eötvös Experiment in Earth Orbit, Davies, R. W. (Ed.), Proc. Conf. on Experimental Tests of Gravitation Theories **JPL TM**, 33-499, p. 228 (1970); Everitt, C. W. F., Damour, T., Nordtvedt, K., Reinhard, R. Historical perspective on testing the Equivalence Principle, Advances in Space Research **32**, 1297–1300 (2003).
4. Bonneville, R. GEOSTEP: A gravitation experiment in Earth-orbiting satellite to test the Equivalence Principle, Advances in Space Research **32**, 1367–1372 (2003); Sumner, T. J. et al. STEP (satellite test of the equivalence principle), Advances in Space Research **39**, 254–258 (2007).
5. Touboul, P., Willemenot, E., Foulon, B., Josselin, V., Accelerometers for CHAMP, GRACE and GOCE space missions: synergy and evolution, Bollettino di geofisica teorica ed appplicata **40**, 3-4, p. 321-327 (1999).
6. Rodrigues, M. et al. MICROSCOPE Mission scenario, ground segment and data processing. Classical and Quantum Gravity **39**, 204004 (2022)
7. Touboul, P. et al., MICROSCOPE Mission: First Results of a Space Test of the Equivalence Principle, *Phys. Rev. Lett.* **119**, 231101 (2017)
8. Touboul, P. et al Space test of the equivalence principle: first results of the MICROSCOPE mission, Classical and Quantum Gravity **36**, 22 (2019).
9. Touboul, P. et al MICROSCOPE Mission: Final Results of the Test of the Equivalence Principle. *Phys. Rev. Lett.* **129**, 121102 (2022)
10. Touboul, P. et al Result of the MICROSCOPE weak equivalence principle test. Classical and Quantum Gravity **39**, 204009 (2022)
11. Liorzou, F. et al. MICROSCOPE instrument description and validation. Classical and Quantum Gravity **39**, 204002 (2022)
12. Touboul, P. et al MICROSCOPE mission analysis, requirements and expected perfor-

mance. Classical and Quantum Gravity **39**, 204001 (2022)
13. Chhun, R. et al MICROSCOPE instrument in-flight characterization. Classical and Quantum Gravity **39**, 204005 (2022)
14. Rodrigues, M. et al MICROSCOPE: systematic errors. Classical and Quantum Gravit **39**, 204006 (2022)
15. Bergé, J. et al MICROSCOPE mission: statistics and impact of glitches on the test of the weak equivalence principle. Classical and Quantum Gravity **39**, 204008 (2022)
16. Baghi, A., Métris, G., Bergé, J., Christophe, B., Touboul, P., Rodrigues, M., Gaussian regression and power spectral density estimation with missing data: The MICROSCOPE space mission as a case study, *Phys. Rev.* D **93**, 12, p. 122-007 (2016)
17. Bergé, J. et al MICROSCOPE mission: data analysis principle. Classical and Quantum Gravity **39**, 204007 (2022)
18. Bergé, J. MICROSCOPE's view at gravitation. Reports on Progress in Physics **86**, 066901 (2023)
19. Bergé, J., Brax, P., Métris, G., Pernot-Borràs, M., Touboul, P., Uzan, J.-P. MICROSCOPE Mission: First Constraints on the Violation of the Weak Equivalence Principle by a Light Scalar Dilaton. *Phys. Rev. Lett.* **120**, 141101 (2018)
20. Bergé, J. et al MICROSCOPE's constraint on a short-range fifth force. Classical and Quantum Gravity **39**, 204010 (2022); Pernot-Borràs, M. et al Constraints on chameleon gravity from the measurement of the electrostatic stiffness of the MICROSCOPE mission accelerometers. *Phys. Rev.* D **103**, 064070 (2021)
21. Fischbach, E., Gruenwald, J. T., Krause, D. E., McDuffie, M. H., Mueterthies, M. J., Scarlett, C. Y. Significance of composition-dependent effects in fifth-force searches. Physics Letters A **399**, 127300 (2021)
22. Willemenot, E., Touboul, P. On-ground investigation of space accelerometers noise with an electrostatic torsion pendulum. Review of Scientific Instruments **71**, 302–309 (2000)
23. Armano, M. et al Beyond the Required LISA Free-Fall Performance: New LISA Pathfinder Results down to 20 $\mu$ Hz. *Phys. Rev. Lett.* **120**, 061101 (2018)
24. Armano, M. et al Sub-Femto-g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results. *Phys. Rev. Lett.* **116**, 231101 (2016)
25. Abrykosov, P. et al Impact of a novel hybrid accelerometer on satellite gravimetry performance. Advances in Space Research **63**, 3235–3248 (2019)
26. Marque, J.-P., Christophe, B., Foulon, B. Accelerometers of the GOCE Mission: Return of Experience from One Year of In-Orbit. ESA Living Planet Symposium **686**, (2010)
27. Speake, C. C. Forces and force gradients due to patch fields and contact-potential differences. Classical and Quantum Gravity **13**, A291–A297 (1996)
28. Nobili, A. M., Anselmi, A. Relevance of the weak equivalence principle and experiments to test it: Lessons from the past and improvements expected in space. Physics Letters A **382**, 2205–2218 (2018)