On the evaluation of GG by the Science Assessment Review Panel (SARP)

Anna M. Nobili (on behalf of the GG team)

June 2015

In its final *Evaluation of "Galileo Galilei" proposal* among those shortlisted for the M4 competition the *Science Assessment Review Panel (SARP)* appointed by ESA makes the following specific criticisms:

1. SARP states that the GG experiment may be affected by the fact that the test masses are spinning in the gravitational field of Earth, which is also spinning. This criticism was expressed in the file: *Questions to the "Galileo Galilei" Team* received from SARP on April 10, 2015 during the evaluation process of the 10 proposals shortlisted for the M4 competition. The answer we provided (*"Galileo Galilei-GG" Answers to Questions as formulated by ESA M4 "Science Assessment Review Panel"-SARP*, p.4) showed quantitatively that all such effects are smaller than the differential target signal of GG by orders of magnitude. No further question was posed on this issue by SARP and SSC-M4 panel members during the GG interview at ESTEC on April 21. Nonetheless, the final SARP report maintains this criticism (though no indication is given on the dangerous effect and its order of magnitude). Since most UFF/WEP experiments use rotation to up-convert the signal frequency, the issue is of wide interest and we will soon submit for publication a paper on the subject. Individual members of SARP will thus have the possibility to make their objections in writing, so that a firm conclusion can be reached.

2. SARP states that "in some circumstances" (of which no detail is given) whirl motions can become chaotic. In the file: *Questions to the "Galileo Galilei" Team*, SARP referred to the presence of bearings (between a rotor and a stator). This means that the reviewing panel evaluated GG under the assumption that –like all ground rotors– it needs a motor and bearings. However, this is a serious misconception because in space, after initial spin up to the nominal frequency of 1 Hz, GG continues to spin by conservation of angular momentum with no motor, no bearings and no stator! Therefore, all noise sources related to them (including some examples of chaos reported in the literature in the presence of rotating and non-rotating components) are absent in GG. Chaos is also known to occur in the presence of high non linearities (multiple attractors) and large disturbances. This is by far not the case in GG where the largest relative displacements, due to Earth tides coupled to whirl motions (at the whirl frequency, far away from the signal), are of a few hundreds of picometers and the linear approximation always holds. Moreover, small divisors, which are known to generate chaotic motions, are not present in GG because the frequencies involved (spin, whirl and orbital frequency) are far away from each other or away from small exact ratios. Finally, whirl motions in GG are damped every few days; therefore, there is no danger of exceeding the "predictability horizon" beyond which a system may become chaotic (over long timescales even planets are chaotic). If SARP reviewers had explicitly reported (with reference to the literature) the "circumstances" they refer to, it would have been possible to establish if they occur in GG and if so, in which quantitative conditions (e.g. strength of perturbations and non linearities) they may generate whirl chaos over the short timescales of interest.

3. SARP states that for the GG test to be credible it should have an additional accelerometer with equal composition test masses as a check.

   The *GG Proposal* reports the checks of systematics errors that will be carried carry out. The best torsion balance experiments report their analyses of systematics and no test with equal composition masses. The same (obviously) occurs for Lunar Laser Ranging tests. Nonetheless, all these results are accepted in the scientific community on the basis of their respective analyses of systematics. Of course, in order to check systematics it is mandatory to perform a sufficient number of measurements in conditions such that the most relevant disturbances can be discriminated from the signal. In GG this is certainly possible because of the short integration time and the fact that the dynamical conditions change from one measurement to the next because the spin axis is fixed in space while the orbit plane (where the signal acts) changes due to the nodal regression caused by the quadrupole moment of the Earth. In GG all this occurs in a totally passive manner because the satellite is spin-axis stabilized, while in the case of Microscope it requires active control, which may disturb the check. If only one or two measurements are performed in the total duration of the mission, obviously no check of systematics is possible. This was the case of the STE-QUEST proposal –studied for several years as ESA M3– which did not have an equal composition test either. A check with equal composition test masses has been performed only in the mass dropping test with a vertical disk (PRL 69, 1722, 1992) because in that case it was crucial to assess that the test was limited by release errors in dropping the disk at initial time. It is well known that initial condition errors apply only to mass dropping tests and not to tests such as torsion balances and GG.

   In stressing the need for an additional accelerometer with equal composition masses SARP disregards two well known facts: that the two accelerometers cannot necessarily be exactly identical (because test masses in the two accelerometers are obviously not identical); that the spacecraft has only one center of mass and therefore –unless the two accelerometers are concentric, as they are not in Microscope– they are bound to be subjected to different disturbances.

More in general, SARP states that: a) there is no believable mathematical description of the GG dynamics in 3D; b) the description of the instrument and of the experiment is very vague; c) it is not clear that all noise sources have been quantitatively accounted for; d) the proposal is contradictory on the laser gauge read-out.

With a 51 page limit imposed by ESA there could not possibly be enough room to present these issues in all their aspects. We inevitably had to rely on the fact that reviewers would refer to the literature (key features of the GG experiment and its prototype are published in highly routed journals) and to reports of industrial studies available on the web. The latest and most comprehensive study of GG resulted in some 30 documents which, as stated in the Executive Summary, were available on request.

On April 21, during the GG presentation to ESA panels, it was complained by one SARP member that no GG simulation in 3D is available. Indeed, a full simulation of the GG experiment in 3D has been performed in 2009 based on the GOCE simulator (which the GOCE mission proved to be totally successful) and in realistic physical conditions. The existence of the 3D simulation was mentioned in the proposal and its main features were summarized in the *Appendix: the GG simulator* which we provided in response to the *Questions to the "Galileo Galilei" Team.* Since the Appendix may have been incomplete, the relevant original documents of the industrial study were handed over to the panel just after the presentation. In the final assessment SARP maintained its objections on the mathematical model used for GG. However, it has not provided an indication, albeit brief, of what is wrong with it. Since the basic equations have been published also in scientific journals, it should be possible to show in writing where are the errors.

Similarly, SARP members are convinced that not all noise sources have been taken into account in the error budget, but do not mention even one such effect. Some disturbances of which we reported the expected value in the tables and in the error budget (there was no room to present the full derivation for each of them) are generically criticised but no indication is given of which specific one is wrong and by how much.

A laser gauge with a noise level of $1\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ in 1 s was a considerable achievement several years ago when Mike Shao made it at JPL for SIM (and later proposed for GG), but nowadays it is no longer a big challenge. Although a specific one for GG has not been developed (in the prototype we use capacitors), it could be done to TRL5 during the definition study (a possibility specifically allowed by the M4 Call). The great advantages of the laser gauge over the capacitance read-out are very clearly explained in the GG proposal, as well as in the *GG interview slides* presented to SARP and SSC-M4 panels: it allows large gaps (hence reducing gas damping noise and yielding a very short integration time) and it is highly differential (i.e. very good for common mode rejection). Since the GG target signal is about 1 pm, to be measured in a few hours (as clearly stated in the first table of the GG proposal on "Driver # 1: Signal, integration time and readout noise") it is obvious that also a 10 or even 50 times more noisy laser gauge could make it. However, if $1\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ in 1 s is feasible, why should we not aim at it? Also in view of improving the GG test by one order of magnitude, to 10^{-18}. Here it is explained the "contradiction" that SARP refers to: the laser gauge is crucial and yet it is not the limiting factor. As for the noise sources, since the target is not a big challenge and all necessary expertise is available within the core team members of the GG collaboration (in addition to help from Mike Shao) in order to reach TRL5, we did not use too much space in the proposal on this. However, the slides presented at the interview contained positive results on the noise effects pointed out by SARP, and these results have not been questioned during the interview. No specific noise source is mentioned in the final report such that it would prevent the laser gauge from reaching the target noise level.