On the universality of free fall, the equivalence principle, and the gravitational redshift

A. M. Nobili
Department of Physics "E. Fermi," University of Pisa, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy and INFN-Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

D. M. Lucchesi
INFN-Istituto Nazionale di Fisica Nucleare, Sezione di Pisa, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy and INAF-IAPS-Istituto di Astrofisica e Planetologia Spaziali, Via Fosso del Cavaliere 100, 00133 Roma, Italy

M. T. Crosta
INAF-Osservatorio Astrofisico di Torino, Pino Torinese, Torino, Italy

M. Shao and S. G. Turyshev
Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, California 91109

R. Peron
INAF-IAPS-Istituto di Astrofisica e Planetologia Spaziali, Via Fosso del Cavaliere 100, 00133 Roma, Italy

G. Catastini and A. Anselmi
Thales Alenia Space Italia, Strada Antica di Collegno 253, 10146 Torino, Italy

G. Zavattini
Department of Physics, University of Ferrara, Via Saragat 1, 44122 Ferrara, Italy and INFN-Istituto Nazionale di Fisica Nucleare, Sezione di Ferrara, Via Saragat 1, 44122 Ferrara, Italy

(Received 10 August 2012; accepted 15 March 2013)

Through the contributions of Galileo, Newton, and Einstein, we recall the universality of free fall (UFF), the weak equivalence principle (WEP), and the strong equivalence principle (SEP), in order to stress that general relativity requires all test masses to be equally accelerated in a gravitational field; that is, it requires UFF and WEP to hold. The possibility of testing this crucial fact with null, highly sensitive experiments makes these the most powerful tests of the theory. Following Schiff, we derive the gravitational redshift from the WEP and special relativity and show that, as long as clocks are affected by a gravitating body like normal matter, measurement of the redshift is a test of UFF/WEP but cannot compete with direct null tests. A new measurement of the gravitational redshift based on free-falling cold atoms and an absolute gravimeter is not competitive either. Finally, we compare UFF/WEP experiments using macroscopic masses as test bodies in one case and cold atoms in the other. We conclude that there is no difference in the nature of the test and that the merit of any such experiment rests on the accuracy it can achieve and on the physical differences between the elements it can test, macroscopic proof masses being superior in both respects.

[http://dx.doi.org/10.1119/1.4798583]

I. INTRODUCTION

The universality of free fall (UFF)—that the acceleration imparted to a body by a gravitational field is independent of the nature of the body—was established by Galileo and Newton as an experimental fact, within the limits of the experiments of their time. If it holds, inertial and gravitational mass are equivalent; this is the weak equivalence principle (WEP). Newton made this "principle of equivalence" the basis of classical mechanics. More than two hundred years later, Einstein extended it to the invariance of physical laws in nonrotating laboratories freely falling in a uniform gravitational field, making it the foundation of his theory of general relativity (GR).

Because UFF is an experimental fact, it can be disproved by experiment. For Einstein, the clever torsion balance experiments carried out by Eötvös in Hungary at the turn of the 20th century[1] were proof enough. And there the matter stood for about fifty years. Experiments testing general relativity and its foundation, the UFF, were few and required measuring tiny effects to excruciating precision. General relativity came to be regarded as a theory that was beautiful but uninteresting to experimental physicists. Things began to change in the late 1950s, at the dawn of the space age, when physicists realized that the means to test GR were becoming available in the laboratory and, even more promising, in space.

In 1958, a few months after the launch of the first artificial satellite by the Soviet Union, a committee appointed by U.S. President Eisenhower—whose advice led to the establishment of NASA the same year—wrote a famous pamphlet[2] in which physics was first in the list of scientific objectives soon to be addressed in space. They proposed a space experiment to measure the gravitational redshift, which they regarded as a crucial test of GR (Ref. 2, p. 7): "Physicists are anxious to run one crucial and fairly simple gravity experiment as soon as possible. This experiment will test an important prediction made by Einstein's general theory of relativity, namely that a clock will run faster as the gravitational field around it is reduced."

In January 1960, Leonard Schiff[3] showed that, as long as a gravity field affects clocks and normal matter alike, the gravitational redshift can be derived from WEP and special relativity and as such it is not a crucial test of GR. Measurements of the gravitational redshift are therefore tests of the underlying WEP, but he argued that even with the best atomic clocks of the time they could not compete with the torsion balance tests reported by Eötvös half a century earlier.

Recently, a proposal has been made for a new space mission devoted to measuring the gravitational redshift and testing WEP with cold atoms.[4] A much better ground measurement of the gravitational redshift has been reported, based on free falling cold atoms.[5] A strong scientific debate is ongoing and this has motivated the present work.

This paper is organized as follows. In Sec. II, we go back to UFF and the equivalence principle of Galileo, Newton, and Eötvös to stress the advantages of null experiments. In Sec. III, we recall how UFF and Newton's equivalence principle (the "weak equivalence principle") led Einstein to the strong equivalence principle (or "Einstein equivalence principle") and to the formulation of general relativity, so that a violation of UFF/WEP would either require that GR be amended or call for a new force of nature. This has led to the quest for UFF/WEP null experiments that are as accurate and precise as possible.

In Sec. IV, following Schiff,[3] we derive the gravitational redshift from the WEP and special relativity and show that—as long as clocks are affected by a gravitating body like normal matter—its measurement is a test of UFF/WEP but it is by far less competitive than direct null tests. A suggested stronger deviation from WEP of clocks as compared to ordinary bodies[6] appears to be beyond the reach of current and planned experiments. The best, so-far controversial, measurement of the gravitational redshift,[5] based on free falling cold atoms in combination with a nearby absolute gravimeter, is a test of the WEP. As such it is in perfect agreement with the original experimental result,[7] but it is in no way competitive with UFF/WEP null tests. In this analysis we frequently step into the "Schiff conjecture" as formulated in 1973 by Thorne *et al.*[8] In consideration of the "vigorous argument" between Schiff and Thorne on this issue, we trace the conjecture back to Schiff's original statement in Ref. 3 and report it. We also report the results of the best experiment to date that has compared the effect of a gravitating body (the Sun) on the rate of clocks of different internal structure and in different locations as the solar potential changes over the year.[9]

Finally, in Sec. V, we compare UFF/WEP tests using macroscopic proof masses versus cold atoms to show that, although the experiments are completely different, there is no difference in the nature of the tests and one should pursue the most promising ones, both in terms of sensitivity and in terms of differences in the physical properties of the atoms being tested.

II. UNIVERSALITY OF FREE FALL AND THE EQUIVALENCE PRINCIPLE FROM GALILEO TO NEWTON

In the *Discorsi*[10] (pp. 128–129; pp. 84–85 of the English edition) Galileo describes his tests of the universality of free fall (UFF) made with two pendulums of different composition. The book was published in Leiden in 1638 when Galileo was 74, almost blind, and under house arrest by order of the Pope, but he had made these experiments in the early 1600s.[11] The accuracy of the test was[12] about 10^{-3}.

In 1687, in the opening paragraph of the *Principia*,[13] Newton wrote: "This quantity that I mean hereafter under the name of…mass…is known by the weight…for it is proportional to the weight as I have found by experiments on pendulums, very accurately made…." If inertial and gravitational mass m_i and m_g are the same for all test bodies regardless of their mass and composition, the equations of motion under the gravitational attraction of a source mass M (e.g., the Earth, assumed for simplicity to be spherically symmetric) state that they all fall with the same acceleration:

$$m_i\ddot{\vec{r}} = -\frac{GMm_g}{r^3}\vec{r}, \text{ so } m_i = m_g \text{ implies that}$$
$$\ddot{\vec{r}} = -\frac{GM}{r^3}\vec{r}. \tag{1}$$

If inertial and gravitational mass are equivalent, UFF holds; should experiments invalidate UFF, they would invalidate the equivalence of inertial and gravitational mass as well.

This was the "equivalence principle" from 1687 until 1907. Note that Eq. (1) holds for any position vector $\vec{r}$, in the vicinity of the source body as well as very far away from it, e.g., it applies also to test bodies on Earth falling in the gravitational field of the Sun (as tested by Dicke and his students in the early 1960s[14]), or in the case of the Earth and the Moon falling towards the Sun (as tested with lunar laser ranging[15]).

Eötvös and collaborators[1] first coupled the test masses by suspending them on a very sensitive torsion balance, and were able to test UFF in the field of the Earth to about 10^{-8}. Dicke's torsion balance experiment was the first UFF test in the field of the Sun (to $\simeq 10^{-11}$), followed by Braginsky and Panov[16] (to $\simeq 10^{-12}$). More recent experiments with rotating torsion balances have tested UFF both in the field of the Sun[17] and in the field of the Earth[18] yielding the best limits to date (see Ref. 19, Table 3): UFF is confirmed to about 10^{-12} in the field of the Sun and to about 10^{-13} in the field of the Earth.

It is worth stressing that UFF experiments can reach high accuracy because they can be performed as *null experiments*. The physical quantity of interest in UFF experiments is the relative acceleration $\Delta a = a_1 - a_2$ of the free falling proof masses, from which the dimensionless Eötvös parameter

$$\eta \equiv \frac{\Delta a}{a} \tag{2}$$

is obtained (here $a = (a_1 + a_2)/2$ is the average free fall acceleration of the masses in the gravity field of the source body). The η parameter quantifies a deviation from UFF. If UFF holds, $\Delta a = 0$ and $\eta = 0$; for a given value of a, the smaller the differential acceleration measured, the smaller the value of η, the more accurate the test.

If the experiment is designed to measure the differential acceleration between the test masses, the experiment signal should be zero in the absence of UFF violation (after classical differential effects have been reduced to below the target). In such null experiments no precise theoretical prediction must be made which the measured signal should be compared to in order to obtain the physical quantity of interest.

Galileo's UFF test with pendulums was most probably the first example of an "almost" null experiment. He measured for how long the masses of the two pendulums—which he had released as equally as he could—would keep in step with each other. Eötvös arranged the masses on a torsion balance and measured the deflection angle of the suspension fiber. In the absence of UFF violation there should be no differential effect, hence no deflection.

III. EINSTEIN'S "HAPPIEST THOUGHT" AND THE STRONG EQUIVALENCE PRINCIPLE

In 1907, referring to his work "On the electrodynamics of moving bodies" published two years earlier[20] in which he presented what is known as the special theory of relativity, Einstein wrote as follows (Ref. 21, Ch. V, "Principle of relativity and gravitation," Sec. 17, "Accelerated reference system and gravitational field"):

> Until now we have applied the principle of relativity—i.e., the assumption that the laws of nature are independent of the state of motion of the reference system—only to *nonaccelerated* reference systems. Is it conceivable that the principle of relativity holds also for systems that are accelerated relative to each other?
>
> …We consider two systems of motion Σ_1 and Σ_2. Suppose Σ_1 is accelerated in the direction of its X axis, and γ is the magnitude (constant in time) of this acceleration. Σ_2 is at rest, but situated in a homogeneous gravitational field, which imparts to all objects an acceleration $-\gamma$ in the direction of the X axis.
>
> As far as we know, the physical laws with respect to Σ_1 do not differ from those with respect to Σ_2; this derives from the fact that all bodies are accelerated alike in the gravitational field. We have therefore no reason to suppose in the present state of our experience that the systems Σ_1 and Σ_2 differ in any way, and will therefore assume in what follows the complete physical equivalence of the gravitational field and the corresponding acceleration of the reference system.

In a 1919 manuscript (Ref. 22, p. 364) Einstein wrote:

> When, in the year 1907, I was working on a summary essay concerning the special theory of relativity for the Jahrbuch fuer Radioaktivitaet und Elektronik, I had to try to modify Newton's theory of gravitation in such a way that it would fit into the theory. Attempts in this direction showed the possibility of carrying out this enterprise, but they did not satisfy me because they had to be supported by hypotheses without physical basis. At that point, there came to me the happiest thought of my life, in the following form:
>
> Just as is the case with the electric field produced by electromagnetic induction, the gravitational field has similarly only a relative existence. *For if one considers an observer in free fall, e.g., from the roof of a house, there exists for him during his fall no gravitational field—at least in his immediate vicinity.*

Einstein recalled the same facts in a speech given at Kyoto University in 1922 entitled "How I created the theory of relativity":[23]

> While I was writing this [i.e., a summary essay on special relativity], I came to realize that all the natural laws except the law of gravity could be discussed within the framework of the special theory of relativity. I wanted to find out the reason for this, but I could not attain this goal easily.
>
> …The breakthrough came suddenly one day. I was sitting on a chair in my patent office in Bern. Suddenly a thought struck me: If a man falls freely, he would not feel his weight. I was taken aback. This simple thought experiment made a big impression on me. This led me to the theory of gravity. I continued my thought: A falling man is accelerated. Then what he feels and judges is happening in the accelerated frame of reference. I decided to extend the theory of relativity to the reference frame with acceleration. I felt that in so doing I could solve the problem of gravity at the same time.
>
> …It took me eight more years until I finally obtained the complete solution.

Thus, the well tested UFF leads to the statement that a frame at rest in a (uniform) gravitational field and a uniformly accelerated frame free from gravitational fields (with uniform linear acceleration equal and opposite to the acceleration produced in the gravitational frame) are equivalent for all physical processes.

This is a crucial leap from Newton's equivalence principle, as stated in Sec. II (now referred to as the weak equivalence principle WEP), to the strong equivalence principle SEP (also referred to as the Einstein equivalence principle EEP[24]). It is apparent that should experiments invalidate UFF (and the WEP), they would invalidate the SEP as well.

In his influential Les Houches lectures, given in 1963 and published in 1964,[25] Robert Dicke stated the strong equivalence principle in such a precise, clear, and simple manner that it is worth using his definition (Ref. 25, p. 4):

> The strong equivalence principle might be defined as the assumption that in a freely falling, nonrotating, laboratory the local laws of physics take on some standard form, including a standard numerical content, independent of the position of the laboratory in space and time. It is of course implicit in this statement that the effects of gradients in the gravitational field strength are negligibly small, i.e., tidal interaction effects are negligible.

A few sentences below, Dicke adds (Ref. 25, p. 5): "It is well known that this interpretation of the equivalence principle, plus the assumption of general covariance is most of what is needed to generate Einstein's general relativity."

The route that led Einstein from SEP to the field equations of gravitation[26] and the general theory of relativity[27] required moving beyond Euclidean geometry. If all accelerated frames are equivalent, then Euclidean geometry cannot hold in all of them: gravity is *locally* replaced by a uniformly accelerated system, hence *globally* there cannot be just one such system. On the other hand, as Einstein put it: "Describing the physical laws without reference to geometry is similar to describing our thought without words. We need words in order to express ourselves."[23] Luckily, the required non-Euclidean geometry and related mathematics were available, after Gauss' original intuitions, thanks to the work of Riemann and other mathematicians.

In summary, GR is founded on SEP, which requires UFF to hold. Should experiments invalidate UFF, either GR must be amended because it is not a fully correct theory of gravity or else we are in the presence of a new, so far unknown, physical interaction. UFF experiments test a fundamental physical principle; moreover, they can reach extremely high accuracy because they can be performed as *null experiments*. This is why they are extremely powerful probes of fundamental physics, worth improving whenever possible.

In 1916 the best experimental tests of UFF/WEP were those by Eötvös[1]—initiated in 1889—and Einstein explicitly recognized their relevance (Ref. 27, Sec. 2):

> …This view is made possible for us by the teaching of experience as to the existence of a field of force, namely the gravitational field, which possesses the remarkable property of imparting the same acceleration to all bodies.*
> *Eötvös has proved experimentally that the gravitational field has this property in great accuracy. [Footnote in the original.]

IV. THE GRAVITATIONAL FREQUENCY SHIFT

In 1960 Schiff,[3] by applying to clocks the WEP as tested for ordinary bodies, pointed out—and demonstrated—that the gravitational redshift can be derived solely from the weak equivalence principle and special relativity, as Einstein did in 1911 (see Ref. 28, pp. 101–103). While Einstein referred to sources of radiation, Schiff—similarly to Tolman (Ref. 29, p. 192)—used clocks.

Schiff derived the gravitational redshift in a simple and straightforward manner, limiting the assumptions required to the very minimum. Concerning experiments to measure it, he concluded: "Terrestrial or satellite experiments that would go beyond supplying corroborative evidence for the equivalence principle and special relativity would be very difficult to perform, and would, for example, require a frequency standard with an accuracy somewhat better than 1 part in 10^{18}."

We re-derive Schiff's result and show that nowadays measurements of the gravitational redshift have no chance of matching the current level of ground based tests of UFF/WEP, and even less so the more accurate ones to be performed in space.

As long as WEP holds for clocks and normal matter alike, the configurations shown as (a) and (b) in Fig. 1 (taken from Ref. 3) are locally equivalent. "Locally" equivalent means that they are equivalent as long as the two clocks A and B are separated (along the lines of force) by a distance h much smaller than their distance from the center of mass of the gravitating body (e.g., the Earth). In this case tidal effects between them are negligible and the gravity field they are immersed in is uniform. The clocks, both at rest in the uniform field of the gravitating body, can then be replaced by the same clocks accelerated with equal and opposite acceleration in empty space, their equations of motion being the same in the two cases. WEP is all that is needed to make the two configurations equivalent to each other. Schiff chose to reason in configuration (b), as follows. An observer in an inertial reference frame has three identical clocks A, B, and C. When they are at rest they all tick with the same period T. Then, clocks A and B are accelerated with the same constant acceleration g, clock A being the leading one, while clock C remains at rest and is used for comparison.

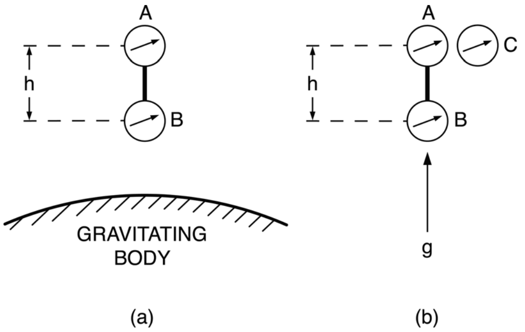

Fig. 1. Illustration from Ref. 3 where the caption reads: "(a) Two identically-constructed clocks, A and B, are at rest in a gravitational field. (b) The gravitating body is replaced by an upward acceleration g of clocks A and B, and a stationary clock C is introduced to compare their periods."

As A and B are accelerated upward with the same acceleration g, clock A will pass by clock C having velocity $v_A > 0$, and therefore it is seen by C as ticking with a slightly longer period T_A with respect to it. Instead, clock B will pass by C—which is still ticking with period T—having a somewhat higher velocity $v_B > v_A$, and therefore ticking with an even longer period T_B. This is because of the time dilation of special relativity, which implies

$$T_A = \frac{T}{\sqrt{1 - v_A^2/c^2}}, \quad T_B = \frac{T}{\sqrt{1 - v_B^2/c^2}} \tag{3}$$

(with c the speed of light). Combining these two equations gives

$$T_B = T_A \frac{\sqrt{1 - v_A^2/c^2}}{\sqrt{1 - v_B^2/c^2}}. \tag{4}$$

By expanding Eq. (4) in power series of v_A^2/c^2 and v_B^2/c^2, we have

$$T_B \simeq T_A \left(1 + \frac{1}{2}\frac{v_B^2 - v_A^2}{c^2}\right) \simeq T_A \left(1 + \frac{gh}{c^2}\right), \tag{5}$$

where we have used $v_B^2 = v_A^2 + 2gh$ because the clocks are moving with uniform acceleration g with respect to the stationary frame of clock C, which is at rest in a gravitation-free region.

The general case of a nonuniform gravitational field (i.e., when h is no longer negligible compared to the distance from the center of mass of the gravitating body) is not substantially different. As Schiff argues, now reasoning in configuration (a), gh can be replaced by the difference $\Delta\Phi$ in the gravitational potential between the positions of clocks A and B. This requires one to perform not just one inter-comparison of clocks but a series of such inter-comparisons between a number of (identical) clocks arranged in such a way that the gravitational field is nearly uniform from one intercomparison to the next. WEP is valid at all distances from the source body, so there will always be a configuration of type (b) where Schiff's arguments hold. Therefore, in a nonuniform gravitational field and as long as second-order terms can be neglected, the periods will be related by

$$T_B = T_A\left(1 + \frac{\Delta\Phi}{c^2}\right) \tag{6}$$

with

$$\Delta\Phi = -\frac{GM}{r_A} + \frac{GM}{r_B} > 0, \tag{7}$$

where M is the mass of the gravitating source body as depicted in Fig. 1(a) (assumed for simplicity to have spherical symmetry), and r_A and r_B are the radial distances of clocks A and B, respectively, from its center of mass. Equation (6) expresses the gravitational redshift $\Delta\nu/\nu = (\nu_A - \nu_B)/\nu_B > 0$ (ν in the denominator indicates the frequency of clock B) stating that the clock at lower altitude (e.g., on the surface of the Earth) is red-shifted if compared to an identical one at higher altitude (e.g., on board of a spacecraft); that is,

$$\frac{\Delta\nu}{\nu} = \frac{\Delta\Phi}{c^2}. \tag{8}$$

It is important to stress that this result has been obtained using neither the conservation of energy nor the mass-energy equivalence.[30] It refers only to identically constructed clocks located at different distances from the center of mass of a gravitating source body along the lines of force. All that is required is that the clocks obey the WEP like ordinary bodies [which makes configurations (a) and (b) of Fig. 1 equivalent to each other] and the special theory of relativity[20] [which provides the time dilation in Eq. (3)].

By Ockham's razor, this demonstration is preferable to others that require more assumptions.

In the experiment we discuss below,[9] the clocks are in free fall relative to the source body, while in GP-A (Gravity Probe A),[31] one clock is in free fall and the other is at rest on the surface of the Earth. As pointed out by Pegna,[32] these facts should not generate confusion as to how the gravitational redshift is derived: the clocks of Fig. 1(a) are at rest relative to the gravitating body, not in free-fall like in an "Einstein elevator." Inside a freely falling elevator, as long as the field is uniform (locally), they would be subject to zero total force, which is equivalent to being inside an elevator at rest in empty space (or moving with uniform velocity), in which case there would be no frequency shift. In Ref. 9, clocks are in free fall together with the Earth relative to the Sun, and the frequency shift occurs due to the variation of the field with the eccentricity of the Earth's orbit. In GP-A,[31] for a clock in free fall near Earth (after separation of the last rocket stage), as the gravitational potential felt by it changes, its frequency relative to the clock at rest is shifted.

We now assume that special relativity is fully valid (as tested by high-energy particle physics) while UFF/WEP might be violated at some level for clocks and ordinary matter alike. The masses of the source body and the clocks are now allowed not to obey the weak equivalence principle of Newton, namely, their inertial and gravitational masses M_i, M_g and m_i, m_g may not be identical. In this case, we write

$$M_g = M_i(1+\eta_e), \quad m_g = m_i(1+\eta_c), \tag{9}$$

where η_e and η_c may differ from zero, quantifying the deviation from equivalence for the source body and the clock, respectively. In the presence of a possible violation of the weak equivalence principle, each clock will be affected by the gravitating body according to the equation (see Ref. 33, Sec. 2)

$$m_i\ddot{\vec{r}} = -\frac{GM_i(1+\eta_e)m_i(1+\eta_c)}{r^3}\vec{r}, \tag{10}$$

and, since UFF/WEP experiments have confirmed the equivalence of gravitational and inertial mass to very small values of η, we can write

$$\ddot{\vec{r}} = -\frac{GM_i[1+\eta+\mathcal{O}(\eta^2)]}{r^3}\vec{r}, \tag{11}$$

with $\eta = \eta_e + \eta_c$. With a separation distance much smaller than the distance from the center-of-mass of the source body, the clocks are subject to a uniform acceleration of magnitude

$$g' = g[1+\eta+\mathcal{O}(\eta^2)], \quad g = \frac{GM_i}{r^2}. \tag{12}$$

If they are separated by a distance comparable to the distance from the source body, they now experience a gravitational potential difference,

$$\Delta\Phi' = \Delta\Phi[1+\eta+\mathcal{O}(\eta^2)]. \tag{13}$$

Equation (8) for the gravitational redshift is correct only as long as $\eta = 0$, which implies $M_g = M_i \equiv M$, $m_g = m_i \equiv m$, $g' = g$, and $\Delta\Phi' = \Delta\Phi$. If $\eta = \eta_e + \eta_c \neq 0$, since the clocks are identical, they are still subject to the same acceleration from the gravitating body, hence configurations (a) and (b) of Fig. 1 are still equivalent to each other and Schiff's method is still correct. However, as shown by Eq. (12), the clocks are subject to an acceleration slightly different than in the case $\eta = 0$ and the gravitational frequency shift involves, to first-order in η, $\Delta\Phi' = \Delta\Phi(1+\eta)$. We therefore have

$$\left(\frac{\Delta\nu}{\nu}\right)_\eta = \frac{\Delta\Phi}{c^2}\left[1+\eta+\mathcal{O}\left(\frac{\Delta\Phi}{c^2}\right)\right], \tag{14}$$

where the order of magnitude of the contribution from gravitational redshift to second-order has been introduced since it may be needed in comparison with η. For the purpose of the comparison, we consider an ideal experiment in which the clocks are at rest relative to each other and the gravitating body is spherically symmetric and non-rotating; in this case all contributions of order $1/c^3$ are zero.[34] As for terms of order $1/c^4$, a detailed derivation performed in the PPN (parametrized post-Newtonian) framework can be found in Ref. 35—including the contribution due to the rotation of the gravitating body. However, only their order of magnitude is relevant here.

Equation (14) shows that the measurement of the gravitational redshift is affected by a WEP violation even if it involves a single set of clocks, all with the same composition (same internal structure). If, in addition, the frequency shift is measured with another set of identical clocks A′, B′, C′ that are different in composition from the set A, B, C, the shift measured will be different if $\eta_{c'} \neq \eta_c$. Assuming that all other experimental conditions are equal, the difference in the frequency shifts measured with the two sets of clocks will depend only on the difference $\eta_c - \eta_{c'}$, while η_e is eliminated and there is no dependence on the composition of the source body. Similarly, in the case of UFF tests that measure the differential acceleration of two proof masses of different composition freely falling in the gravitational field of a source body, only the materials constituting the proof masses are relevant (the experiments test $\eta = \eta_2 - \eta_1$, subscripts 1 and 2 referring to the two proof masses), while the composition of the source body does not affect the measurement.[33]

In summary, as long as clocks are affected by a gravitational field like normal matter, if the possibility of WEP violation is taken into account, the general expression for gravitational redshift is given by Eq. (14). This result shows that a measurement of the gravitational redshift does also test the weak equivalence principle and can therefore be compared with UFF experiments.

Limits on η are set by UFF experiments for the specific materials employed, which are selected with the purpose of maximizing the chance of violation. We regard η measured by UFF as an upper limit for Eq. (14), meaning that if

$$\eta > \mathcal{O}\left(\frac{\Delta\Phi}{c^2}\right), \tag{15}$$

and provided that $\Delta\nu/\nu$ is measured to somewhat better than $\eta\Delta\Phi/c^2$, it follows that the gravitational redshift experiment would test WEP better than UFF experiments. In a real space experiment, because of the motion of the clocks, additional terms would appear in Eq. (14), larger than the one to order $1/c^4$. These terms are of order $1/c^2$ [see Eq. (16)] as well as[34] $1/c^3$, and they all need to be adequately modeled.

Schiff's paper appeared in January 1960. At that time, the best UFF test had been performed by the Eötvös group several decades earlier,[1] to $\eta \simeq 10^{-8}$. The space age had just begun and the GP-A mission proposal, to launch an atomic clock to high altitude and compare it with an identical clock on ground, was already under discussion. In such a situation ($\Delta\Phi/c^2 \simeq 4.3 \cdot 10^{-10}$; this value would slightly increase to $6 \cdot 10^{-10}$ with STE-QUEST[4]), inequality (15) holds and Schiff estimated that a frequency standard with an accuracy somewhat better than one part in 10^{18} was required for a measurement of the gravitational redshift to improve on the UFF tests by Eötvös. Such a good frequency standard was far beyond reach at the time. Hence Schiff concluded that a space experiment to measure the gravitational redshift would have limited significance.

GP-A was launched by NASA 16 years later, in 1976. Two identical hydrogen maser clocks were used to measure the gravitational redshift, one on the ground and the other on board a spacecraft launched to 10,000 km altitude. The result[31] was reported in 1980:

$$\left(\frac{\Delta\nu}{\nu}\right)_{\text{GP-A}} = [1 + (2.5 \pm 70) \times 10^{-6}] \times \left(\frac{\varphi_s - \varphi_e}{c^2} - \frac{|\vec{v}_s - \vec{v}_e|^2}{c^2} - \frac{\vec{r}_{se} \cdot \vec{a}_e}{c^2}\right), \tag{16}$$

where the first term $(\varphi_s - \varphi_e)/c^2$, is the difference in gravitational potential between the clock on the spacecraft and the clock on the ground divided by c^2 (it is our $\Delta\Phi/c^2$).

Two more terms had to be taken into account in order to predict the correct value of the expected first-order gravitational redshift. The second term contains the square of the relative velocity $|\vec{v}_s - \vec{v}_e|$ between the spacecraft and the Earth and it is the second-order Doppler effect (the first-order Doppler effect was nicely canceled in the experiment). The third term is the residual first-order Doppler shift due to the acceleration $\vec{a}_e$ of the ground station during the light traveling time between the spacecraft and the ground station, $|\vec{r}_{se}/c|$.

As shown in Eq. (16), the measured value of $(\Delta\nu/\nu)_{\text{GP-A}}$ agrees with the theoretical prediction to the level 1.4×10^{-4}, which makes η as tested with UFF experiments (to 10^{-12} at the time and to 10^{-13} since 2008) totally inaccessible to GP-A. If the result had shown a discrepancy from the first-order prediction, the experiment itself should have been carefully scrutinized. In this experiment the second-order term of the redshift amounts to $\simeq 1.8 \cdot 10^{-19}$ and in the PPN framework it contains the β and γ parameters[35] ($\beta = \gamma = 1$ in general relativity). That is, to second order the WEP no longer accounts for the gravitational redshift and Schiff's derivation of it is inapplicable, as he noted in his concluding remarks (Ref. 3, p. 343). However, this term is too small to be detected even with a frequency standard of accuracy 10^{-18} and it would not be detected by the STE-QUEST experiment[4] either.

The GP-A measurement of the gravitational redshift shows clearly the difference with respect to a null experiment. The measured frequency shift had to be compared with the sum of the three terms on the right-hand side of Eq. (16), whose values depend on various physical quantities, some of which had to be measured during the experiment itself. It is only by comparing the theoretical prediction and the measured shift that the authors could establish the ratio $[1 + (2.5 \pm 70) \times 10^{-6}]$. No wonder that it took them four years to publish the results of an experiment that lasted only about two hours! The advantage of a null UFF experiment is apparent and we are not surprised to find out that UFF experiments are superior.

Moreover, as clocks become more accurate, the theoretical modeling of all the contributing terms and the measurement to a comparable accuracy of all the physical parameters involved in them become more and more difficult.[34,35]

The accuracy of frequency standards has considerably improved since GP-A. However, since 2008 UFF has been tested to $\eta \simeq 10^{-13}$ in the field of the Earth,[18] hence the η term in Eq. (14) is much smaller than the second-order frequency shift. Even pushing such measurement to second order would—in the words of Schiff—only provide corroborative evidence for the weak equivalence principle. The current values of the β and γ parameters, which are involved in the second-order term, would not be improved either.

In May 1960 Schild[36] analyzed Schiff's derivation of the gravitational redshift, finding that it is correct (though only to first order, as Schiff himself mentioned). He also showed that, unlike the gravitational redshift, light bending cannot be derived purely from the equivalence principle and special relativity as Schiff did in the same paper. In 1968 Rindler[37] too showed that Schiff's procedure is legitimate for clocks but spurious for rods.

Was Schiff correct in applying to clocks the WEP as tested for ordinary bodies?

In 1975, a year before the launch of GP-A, Nordtvedt[6] investigated the gravitational redshift using in addition the conservation of energy. He concluded that WEP violation might affect the clocks used to measure the gravitational redshift more strongly than it would affect the ordinary proof masses used to test UFF, thus invalidating Schiff's derivation of the gravitational redshift. The "amplification factor" would depend upon the specific energy being rearranged in the physical process taking place in the clocks to generate a frequency standard. He quantified the amplification factor as the ratio of the atom rest mass to the mass-energy involved in the frequency generation process. This amounts to saying that the closer the mass-energy involved in frequency generation is to the rest mass-energy of the atom, the closer the clock is to normal matter. For the hydrogen maser clock, assuming that all the mass-energy involved in frequency generation violates WEP, Nordtvedt estimated that it could "amplify" η measured by UFF experiments by about four orders of magnitude. With $\eta \simeq 10^{-12}$ at the time, even a violation four orders of magnitude larger was still too small to be measured by far. Nordtvedt regarded this kind of contribution to gravitational redshift as a test of energy conservation.

Back in 1960, on the next page after Schiff's paper, Dicke[38] argued against it on the grounds that the results of UFF tests cannot be extended to clocks. The dispute appears to have been affected by the ongoing discussion about NASA funding a space mission (GP-A) devoted to testing the gravitational redshift, and the two papers next to each other are very frequently quoted together in the literature. However, it is worth recalling that a few years later, in 1964, Dicke agreed with Schiff. He wrote (Ref. 25, p. 5): "The red shift can be obtained from the null result of the Eötvös experiment, mass energy equivalence, and the conservation of energy in a static gravitational field and static coordinate system." He derived the gravitational redshift from these assumptions, the mass-energy equivalence being confirmed by the Hughes[39] and Drever[40] experiments. A few pages later, while discussing the three famous tests of GR, on the gravitational redshift Dicke wrote (Ref. 25, p. 25): "While this experiment may not be the most important of relativity experiments, it is interesting, and I should like to discuss briefly the experiment of one of my students, Brault, on the redshift of solar lines."

Schiff commented on Dicke's arguments of 1960 with a note added in proof to his paper (Ref. 3, p. 343), which reads:

> The Eötvös experiments show with considerable accuracy that the gravitational and inertial mass of normal matter are equal. This means that the ground state Hamiltonian for this matter appears equally in the inertial mass and in the interaction of this mass with a gravitational field. It would be quite remarkable if this could occur without the entire Hamiltonian being involved in the same way, in which case a clock composed of atoms, whose motions are determined by this Hamiltonian would have its rate affected in the expected manner by a gravitational field. Nevertheless, as stated in the foregoing, I believe that a direct demonstration that the equivalence principle is valid for clocks would be useful. On the other hand, it is evident that experiments of this type could not verify any feature of general relativity theory other than the first-order change in the time scale.

As we understand it, Schiff is arguing that the rates of clocks must be affected by a gravitational field as expected for normal matter, i.e., all in the same way regardless of their internal structure.

Many years later, Schiff's statement has been reformulated by Thorne *et al.*[8] as follows, and ever since named the *Schiff conjecture*: "Any complete and self-consistent gravitation theory that obeys WEP must also, unavoidably, obey EEP." (Einstein equivalence principle[24] is another way of referring to the strong equivalence principle.[25])

Though the Schiff conjecture is very frequently quoted in the literature, we prefer to refer to Schiff's original statement (and to his paper) also in consideration of the fact that a "vigorous argument" took place on this issue between Schiff and Thorne at the 1970 Caltech-JPL conference on experimental tests of gravitational theories, after which Schiff had no chance to go back to the subject in writing because he died just two months later (Ref. 8, p. 3577).

An experiment testing the effect of a gravitational field on the rates of different clocks has been performed by Ashby *et al.*[9] The rate of any clock on the ground is affected by the gravity field of the Sun. Because the solar potential varies over the year due to the eccentricity of the Earth's orbit, the rate of the clock undergoes an annual variation due to the gravitational redshift from the Sun. Should two nearby clocks of different internal structure be affected differently by the gravity field of the Sun, a difference will appear in their annual frequency shifts. Clocks farther apart on the surface of the Earth can also be compared. Over a timespan of seven years, Ashby *et al.*[9] compared the frequencies of four hydrogen masers at NIST (USA) with one cesium fountain clock in the same lab, and also with three more cesium fountain clocks in Europe (in Germany, France, and Italy). The result is that the annual variation of the gravitational potential of the Sun produces on all pairs of clocks the same frequency shift to 1.4×10^{-6}, despite their different structure and also different location on the surface of the Earth. A similar experiment in space[4] must take into account additional terms (to order $1/c^2$ and $1/c^3$) due to the motion of the clocks. Should a discrepancy be found, its interpretation would be very difficult.

In 2010 a measurement of the gravitational redshift four orders of magnitude better than GP-A was reported.[5] It was obtained by re-interpreting the data of a 1999 experiment[7] that measured the absolute value of the local gravitational acceleration g using cold cesium atoms and atom interferometry. In addition, the experiment performed a test of UFF by comparing this value with the value of g as measured by an absolute gravimeter located nearby in the lab in which a laser interferometer monitors the motion of a freely falling corner-cube retroreflector (after correcting for known systematics like tides, polar motion, and others).

The frequency affected by gravitational redshift in this case would be the Compton frequency $\omega_C = m\hbar/c^2$, with m the rest mass of the free falling cold atom (cesium) and $\hbar$ the reduced Planck constant. The gravitational redshift is recovered from the atom interferometry signal—which contains the local gravitational acceleration g—once g has been measured with the absolute gravimeter. The method has been questioned, and a strong dispute is ongoing.[41–46] Here, we wish to stress the following facts. In this measurement the expected first-order frequency shift is extremely small (because $h \simeq 10^{-3}$ m or less); nevertheless, the accuracy is remarkable (7×10^{-9}), better than the goal of the STE-QUEST proposal in space.[4]

As stated by the authors[44] and agreed upon by their opponents,[45] in this case the mass-energy involved is the full mass-energy of the freely falling atom, just as in the previous WEP test,[7] which is indeed the same (and only) experiment performed. Therefore Nordtvedt's[6] amplification factor would be essentially unity. Both the corner-cube reflector and the cesium atoms must obey UFF (see also Sec. V), and therefore a deviation from the predicted gravitational redshift would show a deviation from UFF/WEP according to Eq. (14), the materials involved being cesium and glass.

The measured frequency shift is found to differ from the predicted gravitational redshift by[5] $\beta = (7 \pm 7) \times 10^{-9}$ (note that the authors use the β symbol with no reference to the homonymous parameter in the PPN framework).

Eleven years earlier,[7] with the same experimental apparatus and the same data, the authors stated: "We show that the macroscopic glass object used in this instrument [i.e., the absolute gravimeter] falls with the same acceleration, to within seven parts in 10^9, as a quantum-mechanical caesium atom." After describing the absolute gravimeter and how its run was used to compare with g as measured by the atom inteferometer, they concluded: "A comparison with the value of g we obtained in a two-day run shows a difference of (7 ± 7) p.p.b."

As we can see, the 2010 and 1999 results agree perfectly with each other, but it is apparent that the accuracy of UFF null tests is totally inaccessible to this experiment (7×10^{-9} vs. $\eta \simeq 10^{-13}$ already achieved[18] and $\eta \simeq 10^{-17}$ possible in space[47]).

V. TESTS OF THE UNIVERSALITY OF FREE FALL AND MASS–ENERGY CONTENT OF THE TEST BODIES

The total mass-energy of a body can be expressed as the sum of many terms, corresponding to the energy of all the conceivable interactions and components: $m = \Sigma_k m_k$. The dimensionless Eötvös parameter $\eta = 2[(m_g/m_i)_\alpha - (m_g/m_i)_\beta]/[(m_g/m_i)_\alpha + (m_g/m_i)_\beta]$, which quantifies the violation of equivalence for two bodies of composition α and β, inertial mass m_i, and gravitational mass m_g, is then generalized to

$$\eta_k = \frac{2[(m_g/m_i)_{\alpha_k} - (m_g/m_i)_{\beta_k}]}{(m_g/m_i)_{\alpha_k} + (m_g/m_i)_{\beta_k}}, \tag{17}$$

such that a nonzero value of η_k would define the violation of equivalence between the inertial and gravitational mass-energy of the kth type. The rest mass would contribute to almost all the mass-energy content of the body and therefore to almost all the measured Δa and $\eta \equiv \Delta a/a$.

Dicke's group used aluminum and gold.[14] Dicke investigated their difference in great detail. In Ref. 25 (p. 4) he wrote:

> …gold and aluminum differ from each other rather greatly in several important ways. First, the neutron to proton ratio is quite different in the two elements, varying from 1.08 in aluminum to 1.5 in gold. Second, the electrons in aluminum move with non-relativistic velocities, but in gold the k-shell electrons have a 15 per cent increase in their mass as a result of their relativistic velocities. Third, the electromagnetic negative contribution to the binding energy of the nucleus varies as Z^2 and represents 1/2 per cent of the total mass of a gold atom, whereas it is negligible in aluminum. In similar fashion, the virtual pair field around the gold nucleus would be expected to represent a far bigger contribution to the total energy than in the case of aluminum. Also, the virtual pion field, and other virtual fields, would be expected to be different in the two atoms. We would conclude that in most physical aspects gold and aluminum differ substantially from each other and that the equality of their accelerations represents a very important condition to be satisfied by any theory of gravitation.

Dicke is listing all known different physical properties of an atom of gold as compared to an atom of aluminum. It does not matter at all, in principle, how many of them are used in the experimental test.

Experiments with macroscopic test bodies naturally have at their disposal a very large number of atoms, typically many times Avogadro's number, while in cold atom experiments 10^9 atoms would already be a very large number. The difference is enormous, but if the two experiments test the same atoms, they should confirm or violate UFF/WEP just the same because there is no difference in the mass-energy content of the atoms being tested. The way each experimentalist "manufactures" his/her own "test masses," the way he/she deals with them (by arranging them, manipulating them, etc.), as well as the way their motion is read in order to extract the required signal (i.e., their differential acceleration, which must be zero for UFF to hold), will obviously be completely different in the two cases. Although extremely important for the outcome of the test and the sensitivity it can achieve, these are technicalities.

As pointed out by Pegna,[48] unless evidence is provided that the preparation and manipulation of cold atoms in atom interferometers alters the mass-energy content of the atoms by an amount accessible to the sensitivity of the test, there is no reason to expect a result different from a test with macroscopic bodies of the same composition.

The fact that macroscopic masses and cold atoms fall with the same acceleration in a gravitational field has been proved to 7×10^{-9} (Ref. 7). The experiment compared the absolute value of the free fall gravitational acceleration of cesium cold atoms measured by atom interferometry with the absolute value of the free fall gravitational acceleration of a nearby macroscopic corner-cube reflector measured with laser interferometry. It is apparent that this is not a null experiment, which is definitely a disadvantage.

Since macroscopic test masses have been shown[18] to obey UFF to 10^{-13} (and this is certainly an upper limit), it follows that the experiment of Ref. 7 proves that cesium cold atoms and *any* macroscopic test mass fall with equal gravitational acceleration to 7×10^{-9}. Thus, in particular, cesium cold atoms and a macroscopic body made of cesium would fall with equal gravitational acceleration to this level.

Through a purely theoretical analysis, Storey and Cohen–Tannoudji demonstrated that in a uniform gravitational field the quantum propagator of any object is determined by the action along the classical path.[50] Based on this result, Unnikrishnan argued that the same is true for accelerated objects in free space.[51] Therefore, the outcome of any test of WEP performed in a classical setup will hold in a quantum context.

It is in general very interesting to perform the same test using completely different experiment concepts and experimental apparatus, because systematic errors will be completely different and possibly better understood. The advantage is noticeable if different experiments have comparable accuracy. The weak equivalence principle has been tested with macroscopic test masses, of widely different composition, to an accuracy of 10^{-13}. Experiments with cold atoms have reached 10^{-7} (see Ref. 49); moreover, they have used ^{85}Rb and ^{87}Rb, whose difference in composition by two neutrons limits the physical relevance of the test.

The challenge is for a breakthrough by several orders of magnitude, which would explore a so-far unknown physical domain where chances for a major discovery are higher. The figure of merit for the competing experiments is the accuracy to which they can test UFF between materials as different as possible in their fundamental physical properties. At the present state of the art, such a breakthrough is more likely to occur with macroscopic test masses than with cold atoms.

VI. CONCLUSIONS

In this paper, we have discussed experiments that test the universality of free all and the weak equivalence principle in comparison with experiments that measure the gravitational redshift. Quoting from Einstein's original writings, we have shown how UFF and the weak equivalence principle led him to the strong equivalence principle (or Einstein equivalence principle) and to the formulation of general relativity, so that a violation of UFF would either require that general relativity be amended, or call for a new force of nature. Hence, the need of testing UFF as accurately as possible.

Following Schiff, we have concluded that a measurement of the gravitational redshift is a test of the weak equivalence principle but one that—being an absolute measurement—cannot compete with the direct null tests of the universality of free fall. Nordtvedt's analysis, whereby clocks might violate the weak equivalence principle to a level much stronger than normal matter, implies that a competitive accuracy would still be inaccessible to current and planned redshift experiments.

A recent measurement of the gravitational redshift, performed with data from a previous experiment that used free falling cesium cold atoms in combination with an absolute gravimeter, is a test of the underlying weak equivalence principle. The result is the same as in the original experiment and it is not competitive with WEP null tests.

Finally, we have compared tests of the weak equivalence principle using macroscopic proof masses with those using cold atoms to show that, although the experiments are completely different, there is no difference in the nature of the tests and one should pursue the most promising ones, both in terms of sensitivity and in terms of differences in the physical properties between the atoms being tested. Results obtained with macroscopic bodies are superior in both respects, with prospects for a breakthrough when the experiments will be performed in space.

ACKNOWLEDGMENTS

This work was supported by ASI (Agenzia Spaziale Italiana) and INFN, and it was performed in part at JPL, Caltech, under a contract with NASA. Thanks are due to F. Pegoraro for helping to present this work in the context of the current debate.

[1] R. V. Eötvös, D. Pekár, and E. Fekete, "Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität," Ann. Phys. **373**, 11–66 (1922).

[2] The White House, "Introduction to outer space. An explanatory statement prepared by the President's Science Advisory Committee," available online at <http://www.fas.org/spp/guide/usa/intro1958.html>, 1–15 (1958).

[3] L. I. Schiff, "On experimental tests of the general theory of relativity," Am. J. Phys. **28**, 340–343 (1960).

[4] ESA, "STE-QUEST: Space-time explorer and quantum equivalence principle space test," available online at <http://sci.esa.int/science-e/www/object/doc.cfm?fobjectid=49307>, 1–39 (2011).

[5] H. Müller, A. Peters, and S. Chu, "A precision measurement of the gravitational redshift by the interference of matter waves," Nature **463**, 926–929 (2010).

[6] K. Nordtvedt, "Quantitative relationship between clock gravitational 'redshift' violations and nonuniversality of free-fall rates in nonmetric theories of gravity," Phys. Rev. D **11**, 245–247 (1975).

[7] A. Peters, K. Y. Chung, and S. Chu, "Measurement of gravitational acceleration by dropping atoms," Nature **400**, 849–852 (1999).

[8] K. S. Thorne, D. L. Lee, and A. P. Lightman, "Foundations for a theory of gravitational theories," Phys. Rev. D **7**, 3563–3578 (1973).

[9] N. Ashby, T. P. Heavner, S. R. Jefferts, T. E. Parker, A. G. Radnaev, and Y. O. Dudin, "Testing local position invariance with four cesium-fountain primary frequency standards and four NIST hydrogen masers," Phys. Rev. Lett. **98**, 070802-1-4 (2007).

[10] G. Galilei, *Discorsi e dimostrazioni matematiche intorno a due nuove scienze attenenti alla mecanica & i movimenti locali, in Edizione Nazionale delle Opere di Galilei*, Vol. VIII, p. 128, Barbera Ristampa del 1968, Firenze; *Dialogues Concerning Two New Sciences* (Dover, New York, 1954).

[11] D. Bramanti, G. Catastini, A. M. Nobili, E. Polacco, E. Rossi, and R. Vergara Caffarelli, "Galileo and the universality of free fall," in *Proceedings of the STEP Symposium*, Pisa, Italy 6-8 April 1993, ESA WPP-115, p. 319.

[12] F. Fuligni and V. Iafolla, "Galileo and the principle of equivalence," in *Proceedings of the STEP Symposium*, Pisa Italy 6–8 April 1993, ESA WPP-115, pp. 104–109.

[13] F. Cajori, *Sir Isaac Newton's Mathematical Principles of Natural Philosophy* (University of California, Berkeley, 1934).

[14] P. G. Roll, R. Krotkov, and R. H. Dicke, "The equivalence of passive and gravitational mass," Ann. Phys. **26**, 442–517 (1964).

[15] J. G. Williams, S. G. Turyshev, and D. H. Boggs, "Progress in lunar laser ranging tests of relativistic gravity," Phys. Rev. Lett. **93**, 261101-1–4 (2004).

[16] V. B. Braginsky and V. I. Panov, "Verification of the equivalence of inertial and gravitational mass," Sov. Phys. JEPT **34**, 463–466 (1972).

[17] S. Baeßler, B. R. Heckel, E. G. Adelberger, J. H. Gundlach, U. Schimidt, and H. E. Swanson, "Improved test of the equivalence principle for gravitational self-energy," Phys. Rev. Lett. **83**, 3585–3588 (1999).

[18] S. Schlamminger, K.-Y. Choi, T. A. Wagner, J. H. Gundlach, and E. G. Adelberger, "Test of the equivalence principle using a rotating torsion balance," Phys. Rev. Lett. **100**, 041101-1–4 (2008).

[19] T. D. Wagner "Torsion-balance tests of the weak equivalence principle," Class. Quantum Grav. **29**, 184002-1–15 (2012).

[20] A. Einstein, "Zur Elektrodynamik bewegter Körper," Ann. Phys. **322** (10), 891–921 (1905); "On the electrodynamics of moving bodies," in *The Principle of Relativity. Original Papers* edited by A. Einstein and H. Minkowsky (University of Calcutta, 1920), pp. 1–34, available online at <http://www.archive.org/stream/principleofrelat00eins#page/n7/mode/2up>.

[21] A. Einstein, "Über das Relativitätsprinzip und die aus demselben gezogenen Folgerungen," Jahrbuch für Radioaktivität und Elektronik, **4**, 411–462 (1907); "On the relativity principle and conclusions drawn from it," in H. M. Schwarz, "Einstein's comprehensive 1907 essay on relativity, part III," Am. J. Phys. **45**(10) 899–902 (1977).

[22] G. Holton, *Thematic Origins of Scientific Thought* (Harvard U.P., Cambridge, MA, 1973).

[23] A. Einstein, "How I created the theory of relativity," translated by Y. A. Ono, Phys. Today **35**(8), 45–47 (1982).

[24] C. M. Will, “The confrontation between general relativity and experiment,” Living Rev. Relativity **9**, 1–100 (2006), available online at <http://www.livingreviews.org/lrr-2006-3>.

[25] R. H. Dicke, *The Theoretical Significance of Experimental Relativity* (Blackie and Son, London and Glasgow, 1964).

[26] A. Einstein, “Die Feldgleichungen der Gravitation,” *Preussische Akademie der Wissenschaften*, Sitzungsberichte (part 2), 844–847 (1915); “The field equations of gravitation,” translated by Wikisource <http://en.wikisource.org/wiki/The_Field_Equations_of_Gravitation>.

[27] A. Einstein, “Grundlage der allgemeinen Relativitätstheorie,” Ann. Phys. (ser. 4) **49**, 769–822 (1916); “The foundation of the general theory of relativity,” in *The Principle of Relativity* (Dover, New York, 1952).

[28] A. Einstein, “Über den Einfluss der Schwerkraft auf die Ausbreitung des Lichtes,” Ann. Phys. **35**(10), 898–908 (1911); “On the influence of gravitation on the propagation of light,” in *The Principle of Relativity* (Dover, New York, 1952).

[29] R. C. Tolman, *Relativity, Thermodynamics and Cosmology* (The Clarendon Press, Oxford, 1934).

[30] A. Einstein, “Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?,” Ann. Phys. **18**(13), 639–641 (1905); “Does the inertia of a body depend upon its energy content?” <http://www.fourmilab.ch/etexts/einstein/E_mc2/www/>.

[31] R. F. C. Vessot, M. W. Levine, E. M. Mattison, E. L. Blomberg, T. E. Hoffman, G. U. Nystrom, and B. F. Farrel, “Test of relativistic gravitation with a space-borne hydrogen maser,” Phys. Rev. Lett. **45**, 2081–2084 (1980).

[32] R. Pegna, private communication I (2012).

[33] A. M. Nobili, G. L. Comandi, D. Bramanti, Suresh Doravari, D. M. Lucchesi, and F. Maccarrone, “Limitations to testing the equivalence principle with satellite laser ranging,” Gen. Relativ. Gravit. **40**, 1533–1554 (2008); Gen. Relativ. Gravit. **40**, 1555 (2008).

[34] L. Blanchet, C. Salomon, P. Teyssandier, and P. Wolf, “Relativistic theory for time and frequency transfer to order c^{-3},” Astron. Astrophys. **370**, 320–329 (2001).

[35] T. P. Krisher, “Parametrized post-Newtonian gravitational redshift,” Phys. Rev. D **48**, 4639 (1993).

[36] A. Schild, “Equivalence principle and red-shift measurements,” Am. J. Phys. **28**, 778–780 (1960).

[37] W. Rindler, “Counterexample to the Lenz-Schiff argument,” Am. J. Phys. **36**, 540–544 (1968).

[38] R. H. Dicke, “Eotvos experiment and the gravitational red shift,” Am. J. Phys. **28**, 344–347 (1960).

[39] V. W. Hughes, H. G. Robinson, and V. Beltran-Lopez, “Upper limit for the anisotropy of inertial mass from nuclear resonance experiments,” Phys. Rev. Lett. **4**, 342–344 (1960).

[40] R. W. P. Drever, “A search for anisotropy of inertial mass using a free precession technique,” Philos. Mag. **6**, 683–687 (1961).

[41] P. Wolf, L. Blanchet, C. J. Bordé, S. Reynaud, C. Salomon, and C. Cohen-Tannoudji, “Atom gravimeters and gravitational redshift,” Nature **467**, E1 (2010).

[42] M. A. Hohensee, S. Chu, A. Peters, and H. Müller, “Equivalence principle and gravitational redshift,” Phys. Rev. Lett. **106**, 151102-1–4 (2011).

[43] P. Wolf, L. Blanchet, C. J. Bordé, S. Reynaud, C. Salomon, and C. Cohen-Tannoudji, “Does an atom interferometer test the gravitational redshift at the Compton frequency?,” Class. Quantum Grav. **28**, 145017-1–145017-23 (2011).

[44] M. A. Hohensee, S. Chu, A. Peters, and H. Müller, “Comment on: ‘Does an atom interferometer test the gravitational redshift at the Compton frequency?,’” Class. Quantum Grav. **29**, 048001-1–5 (2012).

[45] P. Wolf, L. Blanchet, C. J. Bordé, S. Reynaud, C. Salomon, and C. Cohen-Tannoudji, “Reply to comment on: ‘Does an atom interferometer test the gravitational redshift at the Compton frequency?,’” Class. Quantum Grav. **29**, 048002-1–3 (2012).

[46] C. S. Unnikrishnan and G. T. Gillies, “Reexamining the roles of gravitational and inertial masses in gravimetry with atom interferometers,” Phys. Lett. A **377**, 60–63 (2012).

[47] A. M. Nobili, M. Shao, R. Pegna, G. Zavattini, S. G. Turyshev, D. M. Lucchesi, A. De Michele, S. Doravari, G. L. Comandi, T. R. Saravanan, F. Palmonari, G. Catastini, and A. Anselmi, “‘Galileo Galilei’ (GG): space test of the weak equivalence principle to 10^{-17} and laboratory demonstrations,” Class. Quantum Grav. **29**, 184011-1–18 (2012).

[48] R. Pegna, private communication II (2012).

[49] S. Fray, C. A. Diez, T. W. Hänsch, and M. Weitz, “Atomic interferometer with amplitude gratings of light and its applications to atom based tests of the equivalence principle,” Phys. Rev. Lett. **93**, 240404-1–4 (2004).

[50] P. Storey and C. Cohen–Tannoudji, “The Feynman path integral approach to atomic interferometry. A tutorial,” J. Phys. II France **4**, 1999–2027 (1994).

[51] C. S. Unnikrishnan, “The Equivalence Principle and Quantum Mechanics: A theme in harmony,” Mod. Phys. Lett. A **17**, 1081–1090 (2002).

Air Pump

This big air pump, with a pump-plate fifteen inches in diameter, stands on a turn of the central staircase of Rockefeller Hall at Case Western Reserve University. It was made by the Boston firm of E.S. Ritchie & Sons, and is listed at $250 in the 1891 catalogue. The rotary nature of the driving effort is unusual; most pumps of this era used a lever action. There is an identical pump at Dartmouth College. (Notes and photograph by Thomas B. Greenslade, Jr., Kenyon College)