Does Information Have Mass?

By LASZLO B. KISH

Department of Electrical and Computer Engineering,
Texas A&M University, College Station, TX 77843-3128 USA

CLAES G. GRANQVIST

Department of Engineering Sciences, The Ångström Laboratory,
Uppsala University, SE-75121 Uppsala, Sweden

electron involved in signal transfer carries one bit of information [2], which makes the corresponding mass to be about 10^{-5} g. The difference between the two mass estimates is partially caused by the low energy efficiency of classical computers [3] (note that the efficiency of quantum computers is even worse [4]).

We will address the fundamental question of minimum mass related to a bit of information from the angles of quantum physics and special relativity. Our results indicate that there are different answers depending on the physical situation, and sometimes the mass can even be negative. We tend to be skeptical about the earlier mass estimations, mentioned above, because our results indicate that the electron's mass does not play a role in any one of them.

Finally, in a seemingly related but actually different and independent context of measured inaccuracies of the gravitation constant, we refer to experiments on the weight transient of memory devices after information writing and erasure.

I. MASS OF THE INFORMATION CARRIER DURING COMMUNICATION

A. Single Electrons as Information Carriers

Electrons have a finite rest mass of $m_e = 9.11 \times 10^{-31}$ kg, which is their mass at zero velocity. This is an

Does information have mass? This question has been asked many times and there are many answers even on the Internet, including on *Yahoo Answers*. Usually the answer is "no." Attempts have been made to assess the physical mass of information by estimating the mass of electrons feeding the power-guzzling computers and devices making up the Internet [1], the result being around 50 g. Other efforts to calculate the mass of information have assumed that each

extremely small number, but it is infinitely larger than zero which is the smallest mass a photon theoretically may have.

B. Single Photons

One cannot store information with photons, but it is possible to communicate information with them, either through an optical fiber or via free space. In an exploration of the fundamental lower limits of the transferred mass related to information, it is, therefore, more proper to use photons than electrons. In an ideal case, a single photon can carry one bit of information, for example, by using the two orthogonal polarization modes to represent the two bit values. According to quantum mechanics, the energy E_p of a photon is given by Planck's formula which states that

$$E_p = hf_p \qquad (1)$$

where $h = 6.63 \times 10^{-34}$ Js is Planck's constant and f_p is the frequency of the photon. The relativistic mass m_p of the photon is given by Einstein's famous equation

$$E_p = m_p c^2 \qquad (2)$$

where $c = 3.00 \times 10^8$ m/s is the speed of light. Equations (1) and (2) then yield that

$$\begin{aligned} m_p &= \frac{h}{c^2} f_p \\ &= 2.21 \times 10^{-42} f_p \ (\text{kg}). \end{aligned} \qquad (3)$$

So if a single optical photon with $E_p \approx 1$ eV $\approx 1.6 \times 10^{-19}$ J carries one bit of information, the mass of the carrier of the communicated bit is $m_p \approx 10^{-36}$ kg, which is about a million times less than the electron's rest mass.

Equation (3) seemingly indicates that the mass of the information carrier can approach zero if one uses photons in the low-frequency limit

where $f_p \to 0$. However, this is not so even if one were able to measure single photons of low frequency, and the reason is twofold: one may neither disregard noise (i.e., the thermal radiation background) nor the uncertainty principle.

C. Photons With Thermal Radiation Background

Each photon mode has two thermodynamic degrees of freedom. In accordance with Boltzmann's energy equipartition theorem, the mean thermal energy E_t in a photon mode is

$$E_t = kT \qquad (4)$$

where k is Boltzmann's constant and T is the absolute temperature of the environment. Equation (4) is valid in the classical physical limit of thermodynamics where

$$hf_p \ll kT \qquad (5)$$

and describes the thermal radiation background, which represents the thermal noise of the electromagnetic field. This noise energy is available for each mode, including the polarization modes. In order to have a sufficiently strong signal and high signal-to-noise ratio (SNR) to achieve a small bit error probability (BEP), the energy representing the bit value must be much greater than the corresponding noise energy, which means that

$$kT = E_t \ll E_p = hf_p. \qquad (6)$$

Equation (6) obviously contradicts (5), which is an indication that single photons cannot be used with high SNR and low BEP in the classical physical limit of thermodynamics. On the other hand, the assumption $f_p \to 0$, which is needed for the photon mass m_p to approach zero, implies the classical limit in (5) for any fixed temperature T. To avoid this contradiction in the

classical limit, one must stay away from using single photons.

Looking now at a number N_p of photons, the requirement of having much greater signal energy than noise energy yields

$$kT \ll N_p E_p = N_p hf_p. \qquad (7)$$

If this is translated to the mass m_w of a bit-carrying classical electromagnetic wave, one gets a frequency-independent result according to

$$m_w \gg \frac{kT}{c^2} \qquad (8)$$

which corresponds to $m_w \approx 4.6 \times 10^{-38}$ kg at room temperature and for zero-decibel SNR.

D. The Uncertainty Principle

Suppose that the bits must be sent with extraordinarily high speed (which is not typical in today's optical communication owing to limitations in light sources and detectors). Then, the energy–time uncertainty principle poses a new limitation which reads

$$\Delta E \Delta t \geq \frac{\hbar}{2} \qquad (9)$$

where the energy uncertainty ΔE is the root mean square (RMS) error in the energy measurement, the "bit duration" Δt is the width of the time window for generating the photon or photon package with the required bit value, and $\hbar = h/2\pi$. Expression (9) can be written $\Delta E \geq \hbar/(2\Delta t)$, and the mass uncertainty of the communicated bit is

$$\Delta m_b \geq \frac{\hbar}{2c^2 \Delta t}. \qquad (10)$$

Expression (10) becomes significant only when the mass uncertainty is larger than the mass related to the

communicated bit, because zero is the low limit of the resultant mass. A bit duration of one picosecond—corresponding to a clock frequency of 1000 GHz in a serial photonic one-bit communication port—yields

$$\Delta m_b \geq 5.9 \times 10^{-40} \text{ kg.} \qquad (11)$$

The latter result indicates that, with a thermal background corresponding to room temperature, the effect of the uncertainty principle would begin to increase the mass of the bit only at bit frequencies exceeding $\sim 10^5$ GHz for the case of a serial one-bit communication line.

Finally, one should note that it is possible to communicate without emitting any signal energy into the communication channel, namely, by modulating the statistical properties of its thermal noise or quantum noise [5]. But even in this case the noise medium carrying the information has a mass.

II. MASS OF STORED INFORMATION

Mass related to information storage is different from the problem studied above, because this medium has mass even without storing any information. The question to address is this: How large is the mass change per bit of the storage medium if one stores information bits there? The same question could have been asked about the photonics-based communication discussed before, but then the mass is independent of the bit value and the information content unless the modulation is so fast that it introduces a signal bandwidth comparable to or greater than the frequency of the photons. We will see next that the answer to the question concerning information storage can even be *negative* mass, depending on the definition of the erased state.

We note in passing that two irrelevant terms sometimes pop up on Internet discussions sites: entropy
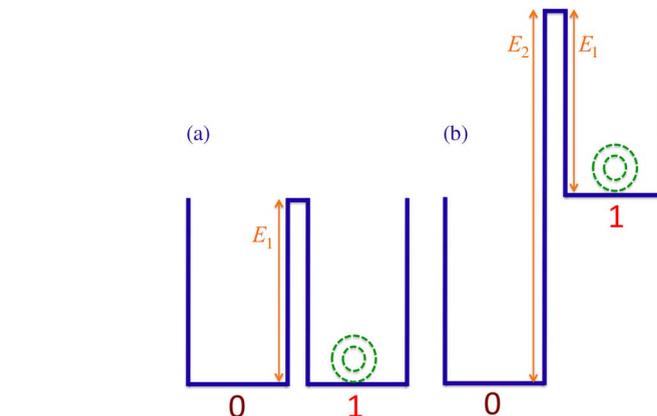
Fig. 1. (a) Symmetric and (b) asymmetric double-well-potential memories. The two bit values correspond to (a) identical and (b) different energies, respectively.

and mass of electrons that charge up a solid-state (flash) memory. A calculation of the thermodynamic entropy does not give any clue to the correct answer, however, because the increase of entropy is relevant for the energy dissipation (which is then released to the environment) rather than for the change of the energy and mass of the system. Similarly, the electron mass does not play any role in these memories because they operate with charge neutrality (i.e., their total charge is always zero, implying that the total number of electrons is constant at any information content). Similarly to any capacitor or metal–oxide–semiconductor (MOS) transistor in a circuit, the charges at the surface of the gate capacitance are always compensated by an equal amount of opposite charges at the other electrode of this capacitor, which is the channel between the drain and the source. Otherwise, a pen drive containing certain information would act like an insulating body charged up to a voltage of 1000–100 000 V; it would attract dust and small pieces of papers and make one's hair stand up when being in its vicinity.

To gain a physical formulation of information storage, we first consider the erased state of the memory. In the case of *normal erasure*, one only alters the address identifying the free memory, and the old information is not

deleted but simply readdressed so that it belongs to the free memory, where it will eventually be written over by new information. Therefore, we discuss *secure erasure* and suppose that 1) all bits have zero value in the (secure-) erased memory; furthermore, we assume that the bits do not interact and ask the question: 2) what is the energy difference between the stored values (0 or 1) of a physical bit?

Fig. 1(a) shows a diagram of the potential energy for a single-bit memory based on a *symmetric* double-well potential. A magnetic memory cell, for example, can approximate this situation. The height E_1 of the potential barrier characterizes the energy that must be invested (dissipated) to change the bit value. Flips corresponding to $0 \rightarrow 1$ and $1 \rightarrow 0$ require the same lower limit of energy E_1.

Fig. 1(b) illustrates a potential energy diagram of a single-bit memory based on an *asymmetric* double-well potential. A floating-gate metal–oxide–semiconductor field-effect transistor (MOSFET), which is the flash-memory element, is an example of this situation (although the energy levels for values 0 and 1 may be inverted). At one of the bit values, the floating gate—which is embedded in an insulating oxide or nitride—must be charged up by a quantum-tunneling current via the insulator in order to keep the channel of the floating-field-effect transistor in the required state.

In Fig. 1(b), the charged position is represented by the high-energy state 1 and the discharged position is represented by the low-energy state 0. The height E_1 of the potential barrier characterizes the energy that must be invested to trigger the bit-value change $0 \rightarrow 1$, and E_2 is needed for the $0 \rightarrow 1$ flip. These potential barriers represent thermal activation in the two directions. (Note that energies that are much larger than E_1 or E_2 are used in practice to facilitate high-speed operation via quantum tunneling, and thus the actual energy dissipation is greater than indicated in our simplified analysis.)

What is the minimum energy barrier E_1 needed in these devices in order for them not to "forget" their information? Supposing that nothing but thermal activation leakage takes place, one gets [6], [7]

$$E_1 > kT \ln\left(\frac{1}{\varepsilon_m}\frac{t_m}{\tau}\right) \qquad (12)$$

where t_m is the lifetime of the memory, $\varepsilon_m \ll 1$ is the probability of bit error during this lifetime, and τ is the correlation time of the thermal fluctuations driving the thermal excitation. Assume that the memory contains N bits and that the goal is that the probability of even a single bit error should be negligible during the memory's entire lifetime. Then, the relation $\varepsilon_m \ll 1/N$ must hold, which leads to

$$E_1 > kT \ln\left(N\frac{t_m}{\tau}\right). \qquad (13)$$

For a one-terabyte solid-state drive with ten years lifetime and $\tau = 10^{-12}$ s, one obtains $E_1 \approx 3.1 \times 10^{-19}$ J at room temperature, which corresponds to ~ 75 kT. Though this energy is dissipated by the writing operation during changing the state, the energy and the mass of the system after the operation remain the same as their original values.

On the other hand, the energy change of the asymmetric system in

Fig. 1(b) is characterized by $\Delta E = E_2 - E_1$, which is the energy needed to keep the MOSFET channel closed (or open, depending on the type of MOSFET). The BEP of that MOSFET channel is controlled by an expression that is analogous to the one for E_1 in (13) [3], and it follows that the energy change ΔE of the system must satisfy

$$\Delta E > kT \ln\left(N\frac{t_m}{\tau}\right). \qquad (14)$$

In conclusion, the minimum mass change of a one-bit flash memory, when flipped into bit value 1, is

$$m_f \approx \frac{kT}{c^2} \ln\left(N\frac{t_m}{\tau}\right) \qquad (15)$$

where Einstein's relation in (2) has been used. This yields $m_f \approx 3.4 \times 10^{-36}$ kg at room temperature. Coincidentally, the absolute value of the mass in (15) is very similar to that of a communicated quantum bit at the chosen conditions. However, the mass in (15) is *negative* if one defines the erased memory element to be a MOSFET with charged gate.

III. EXPERIMENTS: NEGATIVE WEIGHT TRANSIENTS OF MEMORIES

Experiments devised to compare the weight of information storage media before and after recording/erasure have been carried out to 10^{-8} kg accuracy by use of a precision balance, and significant differences—the order of 10^{-6} kg—have been found [8], [9], as discussed below.

The original goal of these experiments was not to compare weights, because the considerations in the section above make it evident that the expected differences are too small to be measurable. Instead the weight transients were discovered during tests [8] inspired by explorations of a hypothetical "fifth force" that conceivably com-

plements electrostatic, strong, weak, and gravitational interactions [10]. The reason behind the quest for a "fifth force" is the extraordinary inaccuracy of the Newtonian gravitational constant G, and that specific materials and environments seem to influence measurements that theoretically should offer excellent accuracy. The "fifth force" hypothesis was forwarded by Fischbach *et al.* in the 1980s but was found untenable. Nevertheless, it stimulated several laboratories to carry out measurements during the 1990s with emphasis on enhancing the accuracy of G [11]. These efforts backfired, however, and the new and supposedly improved measurements yielded even greater inaccuracies than before for the short-range value of G [11]. In 1999, the Committee on Data for Science and Technology (CODATA), under the International Council for Science, decided [11] to increase the uncertainty of G from 128 to 1500 ppm (0.15%). The original $\sim 0.01\%$ accuracy was reclaimed the next year, but even this uncertainty is orders of magnitude beyond that of the acceptable level for other physical constants.

The first author (L. B. Kish) hypothesized that the "fifth force" might be due to some short-range interaction between structures with similar information patterns, shared among the test bodies, and their environment [8], [9]. Such an interaction goes beyond today's physics. Tests were nevertheless executed but no measurable interaction force, at a resolution of 10^{-7}N, was found between two identical 4-GB flash drives filled with identical noise sequences [8], [9].

However, a serendipitous result emerged from these efforts to explore the "fifth force," namely, negative weight transients [8], [9] of the order of 10^{-5} N (corresponding to a weight of about 1 mg) during the recording or secure-erasing information (noises and periodic signals) in the gigabyte range on flash, CD, and DVD media. The relaxation of the weight transient was nearly exponential, with time constants ranging from a few minutes to 15 min [8], [9].

An elevated temperature of the storage medium after recording/erasure could cause a lifting force due to heat convection flow and the ensuing Bernoulli force [8], but the observations were inconsistent with thermal relaxation time constants, which are much shorter than those observed. For example, the weight transients of the DVD disks typically have order-of-magnitude longer time constants than expected for thermal relaxation. Another conceivable explanation for the weight changes is the loss of stored moisture in hygroscopic parts of the media during recording, and a corresponding mass relaxation when moisture is regained. However, the observed dependence of the time constant on the humidity of ambient air does not support this assumption, and neither are other simple explanations satisfactory, such as those involving a changed Archimedes force due to thermal expansion (which is a negligible effect).

We emphasize that the experiments were done carefully to avoid artifacts of shocks to the precision balance during placing/removing memories as well as direct contact of the balance plate with the warm flash drives. Tests against artifacts included repeated placing of the same flash drive without informa-tion recording to ascertain that the values were reproducible and time independent in the absence of information being written into the media.

In conclusion, weight transients observed after writing or secure-erasing information into a memory have not been satisfactorily explained. Experiments with better control of the ambient conditions, for example, by use of vacuum, would be desirable. Finally, we note that the observed effects offer a security application, and weight transients conceivably could serve as a "smoking gun" to evidence that a storage medium has been in recent use. ■

REFERENCES

[1] R. Seitz, *Weighing the web*, ADAMANT, Apr. 9, 2007. [Online]. Available: http://adamant.typepad.com/seitz/2007/04/weighing_the_we.html

[2] S. Cass, "How much does the internet weigh?" *Discover Mag.*, Jun. 2007. [Online]. Available: http://discovermagazine.com/2007/jun/how-much-does-the-internet-weigh

[3] L. B. Kish, "Moore's law and the energy requirement of computing versus performance," *Proc. Inst. Electr. Eng.—Circuits Devices Syst.*, vol. 151, pp. 190–194, 2004.

[4] J. Gea-Banacloche and L. B. Kish, "Future directions in electronic computing and information processing," *Proc. IEEE*, vol. 93, no. 10, pp. 1858–1863, Oct. 2005.

[5] L. B. Kish, "Stealth communication: Zero-power classical communication, zero-quantum quantum communication and environmental-noise communication," *Appl. Phys. Lett.*, vol. 87, 2005, Article 234109.

[6] L. B. Kish and C. G. Granqvist, "Energy requirement of control.," *Europhys. Lett.*, vol. 98, 2012, Article 68001.

[7] L. B. Kish and C. G. Granqvist, "Electrical Maxwell demon and Szilard engine utilizing Johnson noise: Measurement, logic and control," *PLoS ONE*, vol. 7, 2012. [Online]. Available: http://www.plosone.org/article/info:doi/10.1371/journal.pone.0046800

[8] L. B. Kish, "'Gravitational mass' of information?" *Fluct. Noise Lett.*, vol. 7, pp. C51–C68, 2007.

[9] L. B. Kish, "Weight fluctuations of information storage media," presented at the *5th Int. Conf. Unsolved Probl. Noise*, Lyon, France, Jun. 2, 2008. [Online]. Available: http://upon2008.ens-lyon.fr/en/Downloads_files/Kish.pdf

[10] E. Fischbach and C. Talmadge, "6 years of the 5th force," *Nature*, vol. 356, pp. 207–215, 1992.

[11] F. Nolting, J. Schurr, S. Schlamminger, and W. Kundig, "Determination of the gravitational constant G by means of a beam balance," *Europhys. News*, vol. 31, pp. 25–27, 2000.