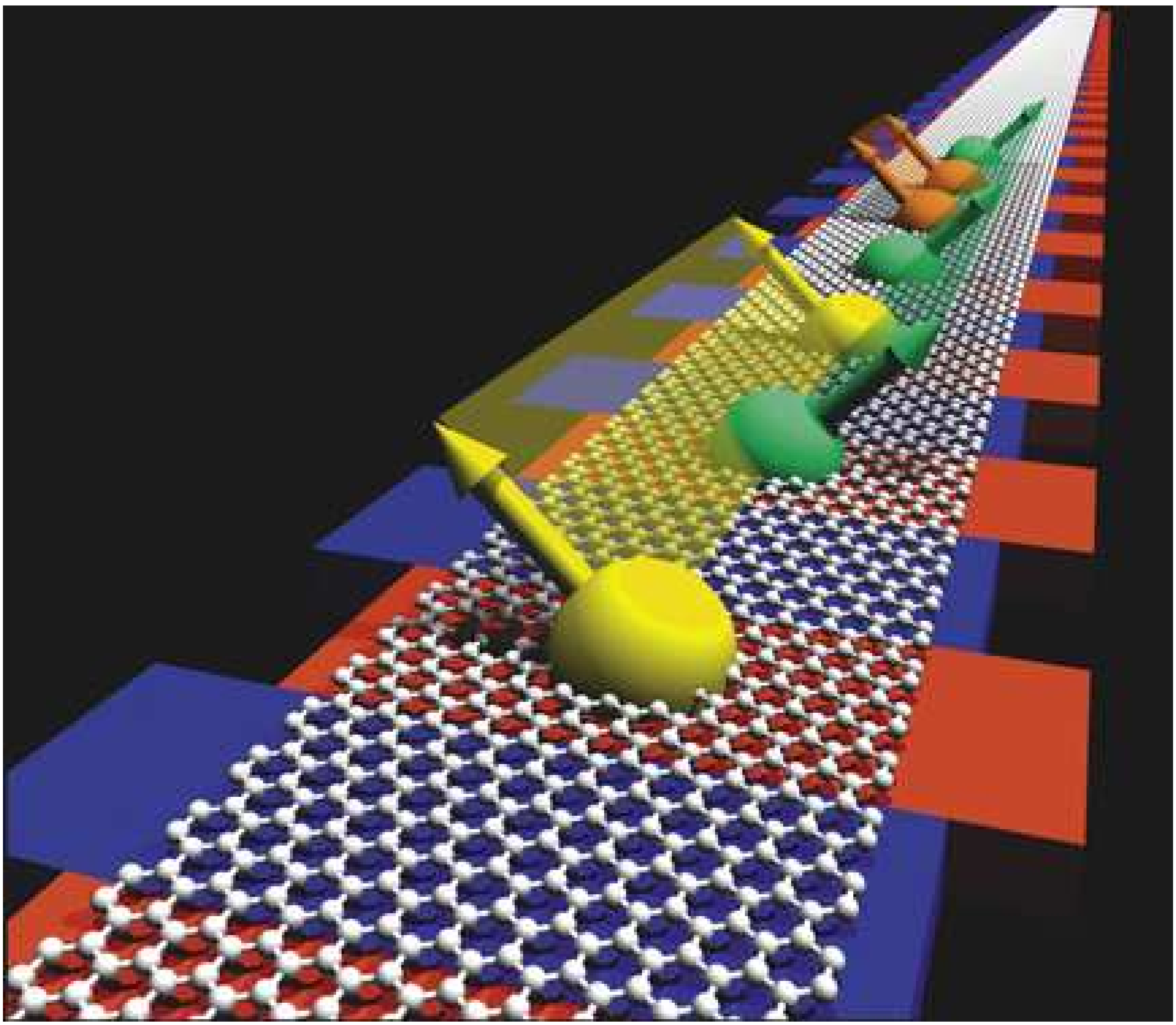
Phys13news

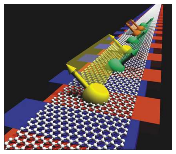
University of Waterloo

Department of Physics & Astronomy
University of Waterloo
Waterloo, Ontario, Canada
N2L 3G1

Winter 2008 Number 125



Spintronics



Cover
Electron spins in graphene (a two-dimensional strip of carbon atoms) could, in principle, be coupled at long range, forming the building blocks for a spintronic quantum computer.
[Image by D. V. Bulaev]

Contents

From the Editor

In this edition of Phys13News we have an article by Bill Coish, a post-doctoral researcher at UW, on spintronics, along with an article by Mike Hudson, a UW faculty member, on dark matter and an article by Rob Bark, an undergraduate student, on electron microscopes. Also, Stefan Scherer returns with another facinating article, this time on the MESSENGER probe and spacecraft navigation.

We also have our usual features, with Rohan Jayasundera, Tony Anderson and George McBernie giving us the SIN Bin, a Sudoku puzzle and prof quotes, respectively.

As always, I look forward to receiving feedback on the content of this issue and suggestions for (or even contributions of) future articles. Please let me know what you like and dislike and topics you would like to see explored in future issues.

Chris O'Donovan is the editor of Phys13news and can be reached at editor2008@phys13news.uwaterloo.ca.

Phys13News is published four times a year by the Department of Physics & Astronomy at the University of Waterloo. Our policy is to publish anything relevant to high school and first-year university physics, or of interest to high school physics teachers and their senior students. Letters, ideas and articles of general interest with respect to physics are welcomed by the editor. You can reach the editor by email at editor2008@phys13news.uwaterloo.ca. Alternatively, you can send correspondence to:

Paper: Phys13News
Department of Physics & Astronomy
University of Waterloo
Waterloo ON N2L 3G1
Fax: 519-746-8115
E-mail: editor2008@phys13news.uwaterloo.ca

Editor: Chris O'Donovan

Publisher: J. McDonnell

Printing: UW Graphics

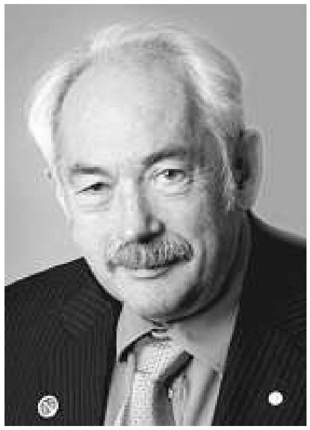



Figure 1: Winners of the 2007 Nobel Prize in physics Albert Fert (left) and Peter Grünberg (right). Photos: U. Montan, Copyright ©The Nobel Foundation, `http://nobelprize.org/`

How to build a better iPod: Spintronics holds the key

W. A. Coish

When the Royal Swedish Academy of Sciences announced the winners of the 2007 Nobel Prize for physics in October of last year, they made it clear that the prize was being awarded for ground-breaking fundamental research that has rapidly led to a new paradigm for the electronics industry [1]. The recipients of the award, Albert Fert of France and Peter Grünberg of Germany (see Fig. 1), independently discovered an effect in 1988 [2,3] known as "Giant Magnetoresistance" (or GMR).

GMR is a phenomenon that converts a very small magnetic moment, arising from the spin degree of freedom of electrons, into a change in resistance to electrical (charge) current. This method of converting spin information into charge–known as a "spin valve"–has been incorporated into the read heads of super-dense hard drives over the last 10 to 15 years. This new technology has made it possible for users to cram 100's of gigabytes of software onto ultra-portable laptops and up to 40 000 of their favourite songs onto an iPod* that can be slipped into a shirt pocket. The benefits do not stop there, according to a growing number of scientists in the nascent field of "spintronics" (short for spin electronics).

One of the great spin-offs of GMR technology was the realization that there was an entire degree of freedom (the electron spin) that had previously been ignored in the development of electronic technology. Traditional electronic devices operate by turning transistors on and off, storing and pushing electron charges around a circuit. Although these techniques have served us well to this point, there are three potential disadvantages to this way of doing things. First, storing charges for a long time is very difficult unless a voltage is continuously applied to keep capacitors charged, so charge memory (what your computer uses for RAM) is usually volatile; it only works when the computer is on. Second, pushing electric charge through resistive elements in a circuit dissipates a lot of power, generating a great deal of heat, with potentially disastrous consequences for expensive hardware [4]. Finally, as circuits are rapidly miniaturized to squeeze more processing power into a smaller package, the rules of quantum mechanics begin to take over; consideration of all relevant microscopic degrees of freedom, including the spin, becomes inevitable.

Proponents of spintronics research hope that spin-based devices may cure many ills of the electronic world. It is well known that magnetic materials are much better suited for the development of robust memory. In fact, tunnelling magnetoresistance (TMR), a close cousin of GMR, is the basis for a spintronic magnetoresistive random-access memory (MRAM), which is now being marketed by Freescale [5]. MRAM is fast and non-volatile, which means that one day it may replace existing hard-drive technology, resulting in a near-instantaneous boot-up. Less energy is required to rotate spins than is typically required to squish electrons through resistors, so many hope that spintronic components would operate more energy efficiently (benefiting the environment) and cooler (benefiting the ambient temperature of your lap). Some spintronic devices show promise to operate much more quickly than their electronic counterparts. Additionally, spin-based devices introduce a new 'knob' to tune the relevant components; they often rely on magnetic-field (rather than electric-field) control. This additional degree of freedom suggests that it may be possible to change physical spin-logic gates on the fly.

There are many fundamental physical questions yet to be answered before new spintronic devices can outstrip conventional electronics in areas other than hard-drive read heads. Some of the most important questions for researchers are: Is it possible to develop semiconducting ferromagnets at room temperature? Can we find simpler ways to combine metallic ferromagnets and semiconductors on the same chip? What are the limitations of "spin injection" (the introduction of spin-polarized electrons into a conducting region)? What limits the lifetime of a spin? and How can one create and control the analogue of electrical currents (spin currents)? Each of these questions has resulted in its own sub-field of research, drawing significant interest from engineers trying to develop clever designs for new components, chemists trying to develop new and exciting materials for nanometre-scale systems, and physicists interested in the inner workings of classical and quantum mechanics (the driving forces behind these devices).

How does it work?

One of the most important fundamental circuit elements in spintronics is the spin valve. A spin valve is a device that allows the passage of one spin orientation, while blocking another. This allows one to create a spin-polarized "beam" of electrons from a typically unpolarized Fermi sea and allows for the readout of spin information through a measurement of electrical current.

One example of a spin valve device is shown schematically in Fig. 2. This device consists of a thin piece of material (shown here in red, which is insulating in the case of TMR and metallic in the case of GMR), sandwiched between two conducting ferromagnetic materials (shown in green). "Up" spins can enter a ferromagnet that is polarized "up" with high probability, whereas "down" spins have a significantly lower probability to enter an "up"-polarized ferromagnet. When both of the ferromagnetic sandwich slices are polarized "up" (top of Fig. 2),

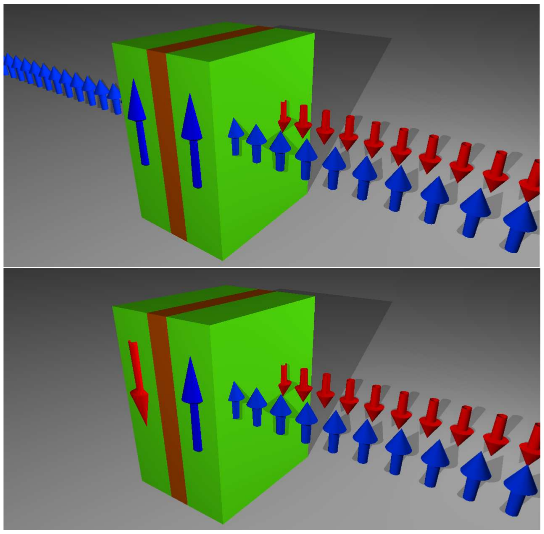


Figure 2: Cartoon of a spin-valve device, similar to those used in hard-drive read heads. Incoming spins with both polarization directions ("up" (blue) and "down" (red)), incident from the right, can only pass with high probability through a ferromagnetic material with similar spin orientation. Only "up" spins pass with high probability when both ferromagnets are oriented in the "up" direction (top) (this is the "open" configuration of the spin valve, allowing current flow). Neither spin orientation can pass through the device when the ferromagnets are oriented in opposite directions (bottom) (this corresponds to the "closed" configuration).

there is significant electronic current through the device due to the flow of "up" spins. In contrast, if one of the ferromagnets flips its spin, current is inhibited for both spin orientations. The presence or absence of current through such a device determines the relative orientation of the two ferromagnets. Such a device is routinely used in hard-disk read heads, where the orientation of one of the ferromagnets is fixed and the other acts as a tiny magnetometer, sensing the 1's and 0's encoded in the magnetic domains of a hard drive.

The spin valve is a convincing demonstration of some of the advantages to using spin in electronic devices (without this device, hard drives would never have reached such small proportions). However, to capitalize on all aspects of the spin, it would be necessary to develop other circuit elements that might channel and manipulate tiny magnetic moments. In a classic paper [6], Datta and Das proposed one such device: a *spin field-effect transistor*, or spin-FET. A conventional electronic FET operates according to the schematic diagrams in Figs. 3(A) and 3(B). When no voltage is applied to the gate electrode (Fig. 3(A)), electrons (blue spheres) travel from the source to the drain along a narrow channel. A voltage applied to the gate electrode (Fig. 3(B)) changes the potential landscape of the channel, preventing current from flowing. In a spin-FET (Figs. 3(C) and 3(D)), the source and drain electrodes are replaced by spin valves (see Fig. 2). If no voltage is applied to the gate electrode in a spin-FET (Fig. 3(C)), spin-polarized electrons retain their polarization as they travel from the source to drain unimpeded, much as unpolarized electrons do in Fig. 3(A). However, when a voltage is applied to the gate of a spin-FET (Fig. 3(D)), the primary effect is not to change the potential landscape of the channel, but rather to enhance spin-orbit coupling. The spin-orbit interaction couples the momentum of the travelling electrons (orbital motion) to the spin degree of freedom. The result of this coupling for electrons moving along the channel is to generate an effective magnetic field (illustrated in Fig. 3(D) with yellow arrows). As the electrons travel down the channel, their spins precess in this effective magnetic field, with the result that the spin may be anti-aligned with the drain valve at the end of the road. If the spin is anti-aligned, with high probability it will not be able to enter the spin valve, and current through the device will be blocked.

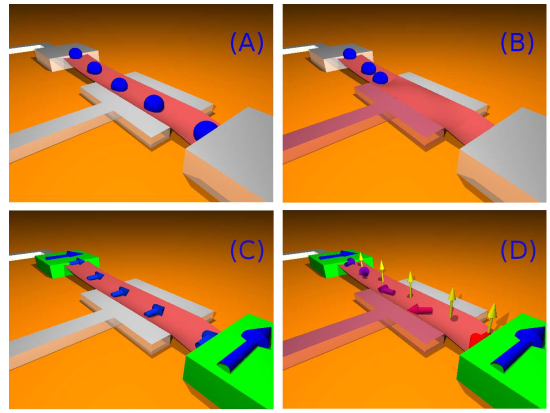


Figure 3: Comparison of a typical field-effect transistor (FET) (Figs. (A) and (B)) with the proposed operation of a Datta-Das spin-FET (Figs. (C) and (D)). When the FET gate is open (Fig. (A)), electrons (blue spheres) flow from the source to the drain through a narrow channel. If a voltage is applied to the FET gate (shown in red, Fig. (B)), the potential in the channel is raised and electron charge can no longer flow from source to drain. In a spin-FET, the source and drain electrodes would be replaced by spin valves (see Fig. 2). When the spin-FET gate is open (Fig. (C)), electrons of a fixed spin orientation emerge from the source and enter the drain, giving rise to a finite current. If a voltage is applied to the gate (Fig. (D)), the electron spins experience an effective magnetic field due to the spin-orbit interaction (yellow arrows). This effective magnetic field causes the spins to precess as they travel from source to drain along the channel. If the spins precess by 180 degrees, they are blocked from entering the spin valve at the drain, resulting in no current flow.

Getting more out of spins

The "conventional" spin-based devices discussed so far (the spin valve and spin-FET) present a significant step forward for computing and electronics. However, harnessing the quantum mechanical properties of these spins could sow the seeds of a revolution. With the full power of quantum mechanics, it would be possible, in theory, to crack codes in a fraction of the time of a classical computer, search databases extremely quickly, send secret messages with guaranteed security, or simulate complicated physical problems with ease. In order to perform many of these tasks, it is first necessary to find some physical realization of an idealized quantum two-level system in which to encode

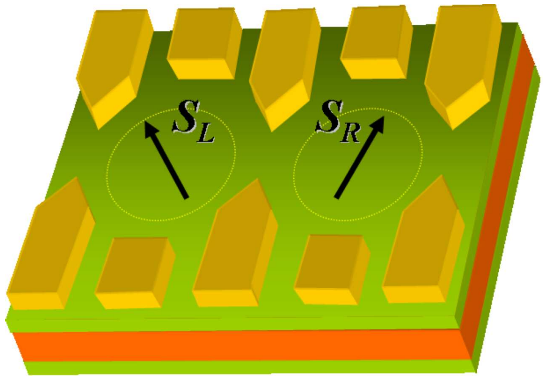


Figure 4: Two electrons confined to a double quantum dot. By applying appropriate voltages to electrodes (shown in yellow), the electrons can be pushed together, increasing the overlap of their wave functions and the exchange coupling between their spins ($\mathbf{S}_L$ and $\mathbf{S}_R$). Changing the voltages to separate the electrons turns the coupling off. Image by V. N. Golovach

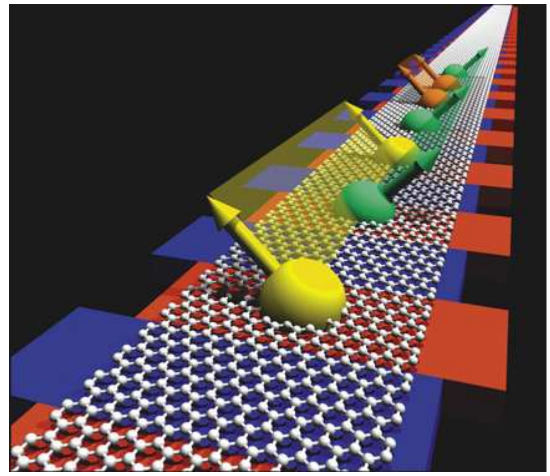


Figure 5: Electron spins in graphene (a two-dimensional strip of carbon atoms) could, in principle, be coupled at long range, forming the building blocks for a spintronic quantum computer. [Image by D. V. Bulaev]

quantum information. These two-level systems, quantum bits or *qubits*, would act as reservoirs of quantum information in a quantum computer, in direct analogy with the bits (zeroes and ones) of a classical computer.

There are many potential qubit-contenders in the race to develop working quantum computers, including trapped atoms and ions, nuclear spins, photon polarization states, superconducting devices, and many others. Since the majority of the electronics industry has developed the art of pushing electrons around in semiconductors in the last 50 years, and the two states of an electron spin-1/2 ("up": $|\uparrow\rangle = |0\rangle$ and "down": $|\downarrow\rangle = |1\rangle$) form the basis for a well-defined quantum two-level system, one natural choice for a qubit is the spin state of a single electron confined in semiconductors. To ensure that only a single two-level system is available for quantum information processing, electrons are often confined to quantum dots (tiny puddles of electron charge that can be designed to hold single electrons – see Fig. 4). Although it is possible to confine single two-level systems constructed from electron spins in such structures, this alone does not yet make a good qubit. The major requirements for any useful qubit can be roughly summarized with two C-words: control and coherence.

In the context of quantum computing, sufficient control is only achieved if two conditions are fulfilled. First, it must be possible to manipulate *single* qubits, independent of each other, with a high degree of accuracy. Second, it is important to be able to perform operations on *pairs* of qubits, typically by selectively coupling and decoupling them. For electrons in quantum dots, single-spin manipulation can be performed using sufficient magnetic field gradients and magnetic resonance techniques. Pairs of electron spins can be coupled selectively through the same strong interaction that generates magnetism: the exchange interaction. Exchange coupling is proportional to the overlap of electron wave functions, and is therefore local and selective for localized electrons. The interaction can be turned on and off between pairs of electrons simply by pushing them closer together or separating them in neighbouring quantum dots (see Fig. 4). One promising extension of this idea suggests that electron spins could be coupled selectively even at long range in materials such as graphene [7] (a two-dimensional sheet of carbon, see Fig. 5).

All the clever coupling schemes in the world do not help to build a quantum computer if, in addition, it is not possible for the qubits to remain coherent for a sufficiently long time. Coherence describes the ability of a qubit to maintain it's identity. That is, if a quantum bit is prepared in a superposition of its two basis states: $\alpha\,|0\rangle + \beta\,|1\rangle$, how long will it remain in this same superposition? This question is one of the most rich and intriguing physical problems associated with quantum computing. On the surface, it also appears to be one of the shortcomings of spintronic quantum computers.

Qubits lose their coherence through interaction with uncontrollable degrees of freedom in the environment. Many physical qubits are designed to minimize this interaction by removing the environment entirely, suspending the qubits in free space. This is the usual approach for trapped atoms and ions. Such isolation does not appear possible for electron spins in semiconductors, where the environment is provided by the semiconducting crystal that holds the electrons in the first place. Instead of removing the environment from spintronic qubits, the strategy of choice is often to control it, by "cooling" the environment to low temperature, judiciously choosing the states of qubits to reduce the environmental influence, or by performing a sequence of quantum mechanical measurements on environmental observables. A new round of experiments is currently underway to determine the viability of these decoherence-minimization techniques. Preliminary results are very promising, but the jury is still out on whether truly useful quantum computation will be possible with such environmental interactions.

Conclusions

As the size scale of conducting channels built into today's transistors reaches a few tens of nanometres, the rules of quantum mechanics become increasingly important, and traditional methods for building faster and cheaper devices will be abandoned. It is not yet clear which new strategies will win out, but some spin-based electronic devices have already taken off. These devices (spin valves in hard-drive read heads and MRAM) may only be the tip of the iceberg.

By exploiting the coherence of quantum spin states fully, it may be possible to build a truly *scalable* quantum computer (one that does not become worse as the number of qubits is increased). With such a device, it should be possible to perform old tasks more quickly, and may give us access to new solutions that were entirely inaccessible with conventional technology.

Further reading

For further reading, see Ref. [8] for a review of GMR, Refs. [9], [10], [11], and [12] for reviews of semiconductor spintronics, and Refs. [13], [14], and [15] for reviews of spintronic quantum computing.

References

[1] Information about the 2007 Nobel Prize in Physics, including video of the Nobel lectures by Albert Fert and Peter Grünberg is available at: `http://nobelprize.org/`.

[2] M. N. Baibich, J. M. Broto, and A. Fert, "Giant Magnetoresistance of (001) Fe/(001) Cr Magnetic Superlattices," *Phys. Rev. Lett*, vol. 61, p. 2472, 1988.

[3] G. Binasch, P. Grünberg, F. Saurenbach, and W. Zinn, "Enhanced magnetoresistance in layered magnetic structures with antiferromagnetic interlayer exchange," *Phys. Rev. B*, vol. 39, no. 7, pp. 4828–4830, 1989.

[4] See a video demonstration at: `http://www.pcpitstop.com/pcsafety/`.

[5] Search for MRAM at: `http://www.freescale.com/`.

[6] S. Datta and B. Das, "Electronic analog of the electro-optic modulator," *Appl. Phys. Lett.*, vol. 56, p. 665, 1990.

[7] B. Trauzettel, D. Bulaev, D. Loss, and G. Burkard, "Spin qubits in graphene quantum dots," *Nature Physics*, vol. 3, pp. 192–196, 2007.

[8] G. Prinz, "Magnetoelectronics," *Science*, vol. 282, no. 5394, pp. 1660–1663, 1998.

[9] S. A. Wolf *et al.*, "Spintronics: A Spin-Based Electronics Vision for the Future," *Science*, vol. 294, no. 5546, pp. 1488–1495, 2001.

[10] D. Awschalom, D. Loss, and N. Samarth, *Semiconductor Spintronics and Quantum Computation*. Springer, 2002.

[11] D. Awschalom, M. E. Flatté, and N. Samarth, "Spintronics," *Scientific American*, June 2002. available online at: `http://www.sciam.com/`.

[12] D. D. Awschalom and M. E. Flatté, "Challenges for semiconductor spintronics," *Nature Physics*, vol. 3, no. 3, pp. 153–159, 2007.

[13] V. Cerletti, W. A. Coish, O. Gywat, and D. Loss, "Recipes for spin-based quantum computing," *Nanotechnology*, vol. 16, no. 4, pp. R27–R49, 2005.

[14] W. A. Coish and D. Loss, "Quantum computing with spins in solids," 2006. Chapter in "The Handbook of Magnetism and Advanced Magnetic Materials", vol. 5 (Wiley). Available at: `http://arxiv.org/abs/cond-mat/0606550`.

[15] R. Hanson *et al.*, "Spins in few-electron quantum dots," *Rev. Mod. Phys.*, vol. 79, no. 4, pp. 1217–1265, 2007.

Bill Coish started in September as a post-doctoral fellow in the Department of Physics and Astronomy and the Institute for Quantum Computing at the University of Waterloo. He completed a B.Sc. (Hons.) in physics from the University of Manitoba in 2000, an M.Sc. from McMaster University in 2002, and a Ph.D. in theoretical physics in 2006 at the University of Basel in Basel, Switzerland under the supervision of Daniel Loss. Bill held a brief (1-year) postdoctoral position in Basel before arriving in Waterloo. He can be reached at wcoish@iqc.ca.

Everything you always wanted to know about Dark Matter ... but were afraid to ask

Mike Hudson

Why should we believe there is dark matter?

Historically, evidence for dark matter came from dynamical measurements. The most well-known are the rotation curves of spiral galaxies, studied by Vera Rubin and co-workers in the late 70s and 80's. The orbits of stars at large radii (in the disk) are believed to be circular. If all the mass were in the centre of the galaxy, one would expect that the orbital velocity would fall as $1/\sqrt{r}$, where r is the distance of a star from the centre of the galaxy, as do the orbital velocities of planets in the Solar System (where essentially all of the mass is in the Sun). Instead, Rubin found that the orbital velocities of stars remained approximately constant as far out in the galaxy as she could measure (see box on page 8). This implies that there is a large amount of mass at large radii, yet there are few stars out there. Hence "dark" matter.

Orbits of galaxies within clusters of galaxies also provide a clue. In fact, as early as 1937 Fritz Zwicky, an astronomer at CalTech, noticed that the velocities of galaxies within the Coma cluster were very fast (typically a thousand km/s). Comparing the total mass derived from the virial theorem to the mass in bright galaxies based on measurements of star light, Zwicky proposed that clusters of galaxies contained large amounts of dark matter, and that this matter could keep the galaxies bound to the cluster. These old results have been confirmed by X-ray observations of clusters. These observations, which measure the pressure of the gas, reveal that there is far more gravitating matter than can be accounted for in observed stars and gas.

Gravitational lensing provides another piece of evidence. Gravitational lensing is a consequence of Einstein's Theory of Relativity. In this phenomenon, a mass (dark or luminous) between the observer and the distant galaxy will bend light rays passing through the cluster, distorting its shape. The amount of distortion can be used to measure the mass that is doing the distorting. By this technique, we also infer far more mass than is present in observed stars or gas. [See issues 116 and 117 for a two-part article by Laura Parker on gravitational lensing. – Ed.]

How much more?

Observations suggest 6-7 times as much gravitating mass as mass observed in stars and gas. So most matter is dark.

Is there any other evidence?

Our understanding of how structure in the Universe – galaxies, clusters and the Cosmic Web of walls and filaments – forms from the first hundred thousand years to the present day also requires there to be dark matter in order to explain the temperature fluctuations in the Cosmic Microwave Background, and the observations of large-scale structure and velocities in the Universe.

Hmmm... OK so what is this Dark Matter?

Most astrophysicists suspect that dark matter is an as-yet undetected particle, one that is fairly massive yet interacts with normal matter only via the weak interaction. Some say that dark matter is a WIMP: a Weakly Interacting Massive Particle.

Why can't the dark matter simply be normal stuff that's dark ... like rocks, planets or black holes?

Two independent cosmological observations – the temperature fluctuations in the Cosmic Microwave Background and the relative abundance of the elements – severely limit the average density of normal matter (matter that can be found in the Periodic Table of the Elements) that can be out there to 3.7×10^{-28} kg/m^3. Yet the measured density of dark matter is some 6-7 times larger ... so while some of the dark matter could be made of normal stuff like planets or bricks, most of the dark matter must be made from stuff that is not found in the Periodic Table of the Elements.

What about black holes?

Black holes are the remnants of massive stars. Since stars are made of stuff from the Periodic Table of the Elements (mostly hydrogen and helium), black holes have to be counted as normal matter. Direct searches for black holes in our Galaxy are not finding evidence that they exist in large enough numbers to make up the dark matter.

I am not sure I buy this. All the evidence for dark matter seems so indirect.

Well, we detect the presence of dark matter through its interaction with the force of gravity. Why should that be less direct than seeing or touching (which, after all, depends on the electromagnetic force)?

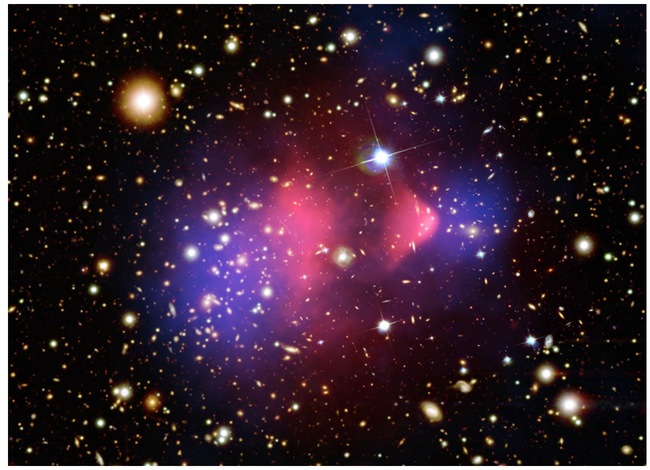


Figure 1: The cluster of galaxies also known as the "bullet cluster," so named because it is thought to be a collision between two galaxy clusters.

An optical image from Magellan and the Hubble Space Telescope shows the galaxies in orange and white.

Hot ionized gas detected in X-rays is seen as two pink clumps in the image and contains most of the "normal," or baryonic, matter in the two clusters. The bullet-shaped clump on the right is the hot gas from one cluster, which passed through the hot gas from the other larger cluster during the collision. The blue areas in this image show where astronomers find most of the mass in the clusters, using the effect gravitational lensing, where light from the distant objects is distorted by intervening matter.

Most of the matter in the clusters (blue) is clearly separate from the normal matter (pink), giving direct evidence that nearly all of the matter in the clusters is dark. The hot gas in each cluster was slowed by a drag force, similar to air resistance, during the collision. In contrast, the dark matter is "collisionless" – so it is not slowed by the impact because it does not interact directly with itself or the gas except through gravity. Therefore, during the collision the dark matter clumps from the two clusters moved ahead of the hot gas, producing the separation of the dark and normal matter seen in the image. If hot gas was the most massive component in the clusters, as proposed by alternative theories of gravity, such an effect would not be seen. Instead, this result shows that dark matter is required. [Image and caption credits: Chandra (NASA)]

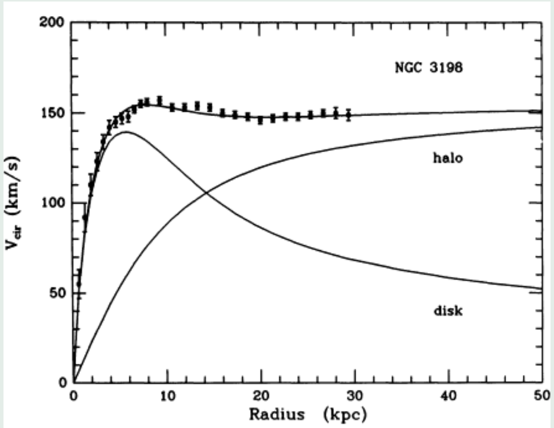
OK fine, but have we ever detected a dark matter particle in the lab?

No, the dark matter particle has not been detected. Several experiments are looking for it, but its a slippery customer. If it interacts only weakly, like the neutrino, then one would expect that it will be hard to detect.

So can this neutrino be the dark matter?

No, the mass of the neutrino is too low. It turns out that the mass of the dark matter particle has an effect on structure formation. If the neutrino had been the dark matter, then the Universe would have been very smooth on a wide range of scales (up to around 100 million light-years). Yet the Universe is observed to be very inhomogeneous on these scales.

Rotation Curves of Spiral Galaxies



The plot shows the orbital velocity of stars and gas as a function of distance from the centre of the galaxy NGC 3198 in units of kiloparcsec (val Albada, Bahcall, Begeman and Sancisi 1985).

The rotation data extend to radii far beyond where most of the stars are. Let's assume that all of the mass is where the stars reside. Then

$$F = \frac{GMm}{r^2} = ma = \frac{mv^2}{r} \quad \Rightarrow \quad v = \sqrt{\frac{GM}{r}}$$

At large radii the velocity should fall as $1/\sqrt{r}$. Instead it seems to be nearly constant as a function of radius.

Recall that if we are inside a spherical object, only the gravity from r applies and the shells outside cause no net force. So we can turn the problem around to find that,

$$M(< r) = \frac{v^2 r}{G},$$

for constant circular velocity. In other words the mass within a sphere of radius r continues to grow linearly with r.

What if our theory of gravity is wrong?

This remains a possibility – after all General Relativity has really only been tested on scales of the Solar System – but its a remote one. Observations of the Cosmic Microwave Background suggest that the normal stuff is reacting to the gravity of dark matter and not just a modified gravity of the baryons. Then there are particular objects, like the well-studied bullet cluster where we see, from gravitational lensing, that the dark matter is segregated from most of the normal matter (Fig. 1).

It remains a possibility that there is theory of gravity which can explain all of these phenomena using just normal matter, but so far no one has succeeded in formulating a convincing one.

What about Dark Energy? What is it? Is it the same as Dark Matter?

No. Dark energy is different. But that's another story...

Mike Hudson is an Associate Professor in the Department of Physics & Astronomy at the University of Waterloo. He can be reached at mjhudson@uwaterloo.ca.

Scanning Tunneling Microscopy

Rob Bark

Scanning tunneling microscopy(STM) involves the direct measurement of surface topography, allowing a more accurate determination of the structure. Images are created as a probe is scanned across the surface of a material with a constant probe tip height or constant tunnel current between the probe tip and the surface (see Fig. 1). Applications of this method yield detailed pictures of a surface's structure in three-dimensions

In 1981, the first experiments involving scanning tunneling microscopy were conducted in Zurich, Switzerland, by Gerd Binning and Heinrich Rohrer. The scanning tunneling microscope (STM) provided a breakthrough in describing the surface

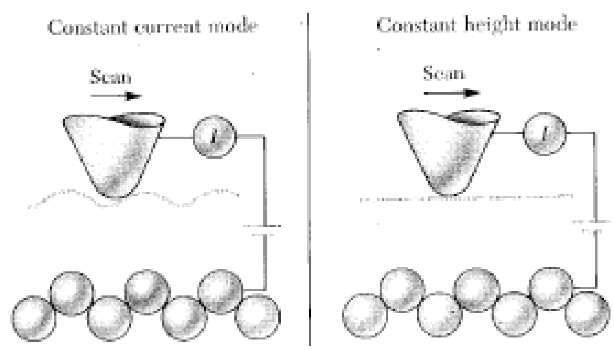


Figure 1: The left side of the figure shows how constant tunneling current measures the surface. Height of the probe changes while scanning. The right figure shows a constant height, implying that the current, I, is changing with x and y position.

structure of materials (including those of non-periodic materials), capturing the 1986 Nobel Prize for Physics. Unlike previous methods, STM allows for a topological, simultaneous scan of the surface structure with a resolution similar to the diameter of a single atom (approximately 1 Å) [1]. What this means for the world of physics, is that it is now possible to gather information about the surface of some material with very high precision without actually making contact with the surface, therefore minimizing, and possibly eliminating, any risk of damage to the sample. Instead of directly providing a visible image to an observer in higher resolution, this microscope creates that image through electrical signals between the probe tip and the material's surface.

The question may arise regarding advantages to STM in place of traditional light or electron microscopes. A light microscope can only detail an image as dictated by the wavelength of light used. This resolution is about 2000 Å at best, and is a far cry from 1 Å, boasted by STM. Electron microscopes have the ability to produce comparable resolution; however, by de Broglie's wavelength ($\lambda = h/p$), the electron speed ($v = p/m$)

required is 1.8×10^6 m/s and at that pace, an electron wouldn't be stopping at the surface. [1]

Instead of a typical picture, STM produces a model that is just as useful and can offer a greater understanding of the material under study.

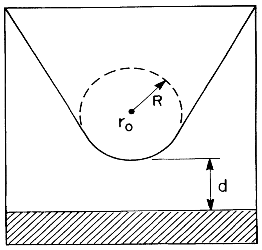


Figure 2: This shows the geometry of the tunneling probe tip. The shape is assumed to be a sphere with its centre at r_0, radius of curvature, R, and distance of closest approach, d.

The region between the probe tip and the surface is known as a vacuum region and acts as an energy barrier, better known as a potential barrier. In classical physics, a potential barrier imposes on an electron the restriction of zero probability to exist in that region. With only classical knowledge, it would be said that an electron cannot move between the surface and the probe tip. Quantum mechanics bridges this gap in the most literal sense. Quantum physics provides an explanation of how a particle can act as though it possesses more energy than it actually does. Small building blocks of matter, such as electrons, exhibit not only properties of particles, but also wave-like attributes that allow electrons to "tunnel," or exist, in areas in which their total energy would not normally permit. Because electrons can tunnel at the quantum mechanical level, some electrons on the surface of the sample metal or on the tip of the probe (depending on the direction of the applied electric field) will travel to the opposing surface. Because the movement of electrons is necessary for tunneling to occur in STM, the material being scanned must have some conductive property.

A unique feature of STM is that it can be performed using two different modes, as distinguished in Fig. 1. These two methods [6] are easily differentiated as constant current versus constant height. To maintain a constant tunneling current, the probe tip adjusts its distance above the surface, and the changing tip height is then collected as an electrical signal. Plotting the signal as a function of the position yields an accurate description of the surface. In contrast to constant current, constant probe height is as its name states. The probe stays at a fixed level and the varying current is collected electronically as a signal, describing the surface under study. The signal, this time recorded as a fluctuation in current, is plotted against the x and y coordinates, giving shape to the surface.

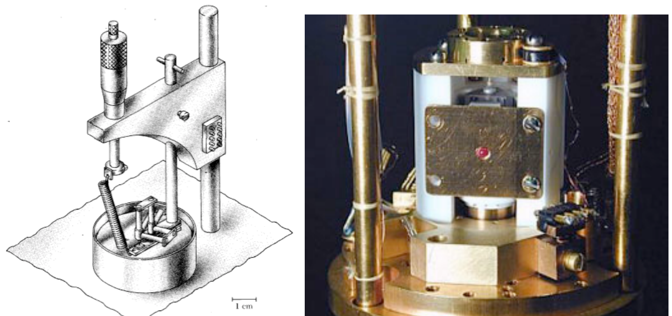


Figure 3: Left is an example of STM equipment and details how the sample and probe tip would be positioned.6 The right hand side pictures an STM used by Laboratory of Atomic and Solid State Physics (LASSP) at Cornell University, New York. [?]

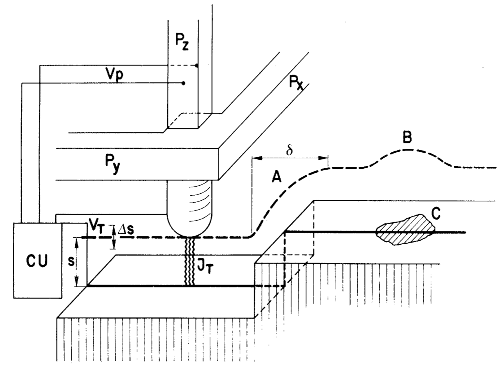


Figure 4: The scanning tunneling microscope. (Distances and measurements are not to scale.) P_x, P_y, and P_z are piezodrives used to scan the surface of the metal under examination. P_x and P_y move the probe tip across the surface of the material, while a voltage (Vp) is applied to P_z by the control unit (labeled as CU) to keep a constant tunneling current, JT, as well as constant tunneling voltage, V_T. The dashed line depicts the probes displacement in the z-direction, scanning along the y-direction (x unchanging). At point A, the probe experiences a step with length δ, and at point B, the surface has a contaminated region, C, having a lower work function, ϕ.

The first use of STM, conducted by Binning *et al.* [2] in the early 1980s, involved a constant tunnel current flowing between a tip formed from Tungsten (W), to the surface of Platinum (Pt), separated by an approximate distance of 10 Å. Constant current allows for the scanning of surfaces that consist of different atomic layers and levels, but sampling is time consuming. In comparison, constant height offers a rapid scanning. In fact, the speed of scanning is approximately 100 times quicker than using constant tunneling current! Since current is measured electronically, there is only lateral movement. In order to create vertical movement of the probe, the piezodrives (described later) utilize a feedback network [1]. The feedback network is a necessity in adjusting tip height. This also decreases the sampling rate, as more electronic signals must be passed through the STM and only then can height be adjusted. Ultimately, less movement means faster sampling. If you believe in the phrase "time is money," then you might ask, "Why consider constant tunneling current when I can get the job done faster with constant height?" Fig. 2 provides a view of two different models of STM. From these images, it can be seen that movement in lateral directions while allowing movement in the third spatial direction can be quite complex, promoting another advantage for constant height methods.

The risk in using constant height is the ease with which the probe tip can be crashed into the surface of the sample, damaging both sample and tip. The solution is not as easy as keeping the tip further away from the surface to prevent contact, as separation distance is very important. It is not an option to know when to raise the probe manually, since the surface is the focus of study! The issues regarding tip value during this process will be addressed later.

The main idea of STM is simple, but is a difficult procedure to conduct due to such small, precise measurements of tip size, separation distance between tip and surface, etc. The metal tip is lowered almost to the surface of the desired material, allowing the vacuum tunneling resistance to be both finite and measurable. Using constant tunneling current, as the tip scans the surface in the x- and y-directions, the height above the changing plane of the surface, d, is recorded. This separation, d, allows measurements in the z- direction to be observed simultaneously. This is an important characteristic, as the improbability of experimental repetition will be discussed.

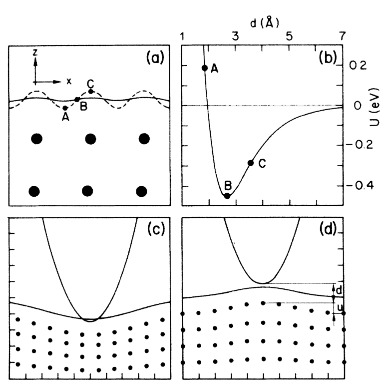


Figure 5: (a) Dashed line is a contour of constant local density of states, where tip-surface force is repulsive at A, reaching a minimum work function at B, and is attractive at C. (b) Potential used during interaction of tip and surface. (c) Compression at point A, and (d) expansion at point C of (a).

The tip of the probe, due to such minuscule size, does not have a well defined structure and thus the geometry and dimensions are forcibly approximated. The lack of consistency in producing the probe tips does not permit exact duplication, but does not mean production is problematic. This implies that any measurements made with one probe tip can only be directly compared to those made by the same tip. This is not a significant problem, ensuring care is taken to make all measurements using one probe. It has also been experimentally proven [3] that results are very similar, even when using other materials for the probe tip (Tungsten, Molybdenum, and stainless steel). From these results, it can be assumed that the tip structure is not critical. The probe tip is speculated as having geometry of spherical configuration with a radius of curvature, R, about a centre r_0. The distance between the probe tip at its closest approach with the surface under examination, can be represented by d, in Fig. 3. The probe tips themselves are only about one or two atoms wide, and although exact duplication is challenging, "sharpening" the tip can be done on a grind stone and has even been accomplished using fine sandpaper – a surprising revelation. [1]

Focusing on the process of utilizing a constant tunnel current, the voltage applied to the piezodrives (the units that scan the surface and convert their measurements to electrical signals) is regulated, as seen in Fig. 4. The current density is defined as the amount of electrical current over a cross-section of the surface area between the tip and surface. The tunnel current is highly dependent on the distance between the probe tip and the surface, as given by current density

$$J_T \propto e^{-A\sqrt{\phi}s}$$

where $A = \sqrt{(4\pi/h)2m} = 1.025$ Å^{-1} eV^{-}, tunnel barrier has an average height (work function), ϕ, barrier separation (also called width), s, m is the free-electron mass and h is Planck's constant. Due to the exponential in this relation, current density has a very sensitive dependency towards small changes in barrier width. Numerically, we can illustrate this sensitivity if we assume a common work function of a few eV - say 3eV. We know A, and let's assume a reasonable separation of 5 Å(0.5 nm). Imagine $JT \approx 1.396 \times 10^{-4}$ in magnitude. With a slight change in separation of 0.01 Å, letting $s = 5.01$ Å, $JT \approx 1.371 \times 10^{-4}$. This small separation change alters the measured current density value approximately 2%! If separation, s, is greater than about 10 atoms (approximately 10 Å), the probability of tunneling is so small that it can be said that tunneling does not occur. On the other side of the coin, STM has the ability, with the appropriate barrier width, to measure surface height on the order of one hundredth of an angstrom. The barrier height, ϕ, is defined as the work function (in electronvolts), and can allow for a change in barrier width of a single surface step (which is 2-5 Å) to cause the tunnel current to differ by three orders of magnitude. The measured length of a step on the surface of the material, observed as δ in Fig. 4, can be approximated by the relation

$$\delta = 3\sqrt{\frac{2R}{A\phi^{1/2}}}$$

where δ is the lateral displacement, as already defined, and A, ϕ, R are defined as above. From Eq. 2, the relationship between the length of a step size and the radius of the probe tip is demonstrated. Intuitively, this is a satisfying conclusion since we would expect a sharper tip to define a step more significantly.

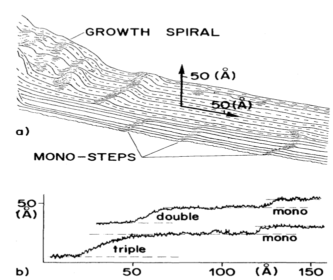


Figure 6: (a) CaIrSn4(110) surface with single atomic steps on the plane and the beginning of a growth spiral as the image shows a rising hill. (b) This shows two separate scans, portraying mono-, double- and triple-atomic steps, with the dashed lines representing the actual (110) faces and their correct measurements.

STM procedures and applications have changed and improved rapidly, allowing for more accurate measurements as a better understanding of vacuum tunneling was developed. Soler et al. [4] conducted studies on STM images dominated by elastic deformations, caused by interatomic forces acting between the surface and the probe tip. From these findings, the relationship between the probe and the surface can be better understood. Results show that corrugations (ridges on the surface of the material) of the image generated by STM increase with tunneling current, while experiencing a constant voltage, V_T. In Fig. 5, the images express how a physical contact between tip and surface is implied, and therefore indicate that tip displacement is not only a factor of electronically induced corrugation. The tip will experience repulsive and attractive forces, as the separating distance depends on not only the tunneling parameters, but also on the local elastic deformation from the tip-to-surface potential and elastic constants of the material. The probe tip was assumed to have the shape of an elliptic paraboloid with radius of curvature, R, at the apex. This estimation of tip geometry appears to be slightly more accurate than the spherical model, and also maintains cylindrical symmetry about the tip axis through the apex, as the previous model does so spherically.

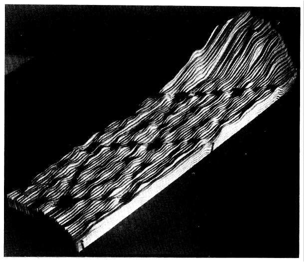


Figure 7: Image of surface topography of Si(111) scanned with STM at 300 Celsius, with a height enhancement of 55% and "hill" in the upper right corner of the picture reaching approximately 15 Å.

Some pictures produced by STM can demonstrate how tunneling reveals the surface details of a material in three-dimensions. Fig. 6(a) shows the topography of a CaIrSn4(110) surface ((110) is the description of the plane and how the specified crystal was cut). Analyzing Fig. 6(b), there are two scans that depict monatomic, double-atomic and triple-atomic steps. These steps are in agreement with theoretical calculations from constant tunneling currents and voltages, similar to the sketch of the step in Fig. 4 at point A.

From other experiments [5] involving STM, illustrating more detailed use of this method of microscopy while studying more complex surfaces of materials, we can begin to directly see the benefits of improved knowledge of surface topography. The surface image of Si(111) in Fig. 7, was taken at 300°C, with the incline on the right of the figure reaching 15 Å. Comparing the resolution of the lower region to the "hill," it can be seen that details collected by STM are precise.

In summary, the effectiveness and accuracy of scanning tunneling microscopy is an important and valuable discovery for the study of surfaces and surface structures.

Using constant current tunneling across a vacuum region between the microscope probe tip and the surface under examination allows detailed imaging of the surface, without imposing any damage. Even with a simple geometric model for the probe tip, the theoretical surface calculations appear to be in close agreement with the experimental picture. Such precise results from approximated calculation appear to hold a promising future, as more detailed understanding of tip structure develops, and STM can be applied to a wider range of experimental areas.

References

[1] Roger A. Freedman and Paul K. Hansma, "Modern Physics" by Serway, Moses, Moyer. Essay: The Scanning Tunneling Microscope, (1997).

[2] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. 49, 57 (1982).

[3] J. Tersoff and D. R. Hamann, Phys. Rev. Lett. 50, 1998 (1983).

[4] J. M. Soler, A. M. Baro, and N. Garcia, Phys. Rev. Lett. 57, 444 (1986).

[5] G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. 50, 120 (1983).

[6] Roger A. Freedman and Paul K. Hansma, "Modern Physics" by Serway, Moses, Moyer. Essay: The Scanning Tunneling Microscope, (1997).

Robert Bark is a third-year undergraduate student in the Department of Physics & Astronomy at the University of Waterloo. He can be reached at rbark@sciborg.uwaterloo.ca.

MESSENGER, Mercury, and General Relativity

Stefan Scherer

"I think Isaac Newton is doing most of the driving now."

Apollo 8 Lunar Module pilot Bill Anders [1], when asked who was driving the capsule on the return from the Moon to the Earth, 26 December 1968.

On January 14, 2008, the MESSENGER spacecraft [2] had a spectacular flyby at Mercury, passing about 200 kilometres above the night-side surface of the planet. While the probe transmitted a wealth of data and an amazing number of exciting photos [3] from this encounter, I would like to focus here on something more ethereal, the influence of General Relativity on MESSENGER's orbit.

As reported on the Planetary Society Weblog [4], the Mercury flyby was a case of "spectacular targeting": after a flight of nearly 100 million kilometers without firing its engines MESSENGER missed the previously planned aimpoint at Mercury by only 1.43 kilometres in altitude. In fact, MESSENGER needs some trajectory fine-tuning from time to time, and the last correction before the flyby, trajectory correction manoeuvre 19 (TCM-19) [5], had occurred 26 days before, on December 19, 2007.

Speaking of a space probe flying to Mercury to a bunch of physics aficionados inevitably brings up General Relativity. After all, the explanation of the extra shift of the perihelion of Mercury [6], a tiny 43 arc seconds per century not accounted for by Newtonian gravitation, was the first big success of General Relativity. So, it seems natural to ask, with such high precision in the determination of MESSENGER's trajectory, how important are effects of General Relativity?

While trying to figure out how the engineers at NASA actually handled gravitation in the trajectory calculations, I realised that a simple back-of-the-envelope calculation already yields a good first estimate for the influence of General Relativity on the space probe. Just applying the formula for the relativistic perihelion shift shows that relativistic effects add up to a few kilometres for the trajectory between the TCM-19 correction and the flyby.

Actually, MESSENGER was, in the months before the flyby, on a quite eccentric elliptic orbit in between the orbits of Venus and Mercury. Fig. 1 is an illustration of part of the orbit from the MESSENGER web site.

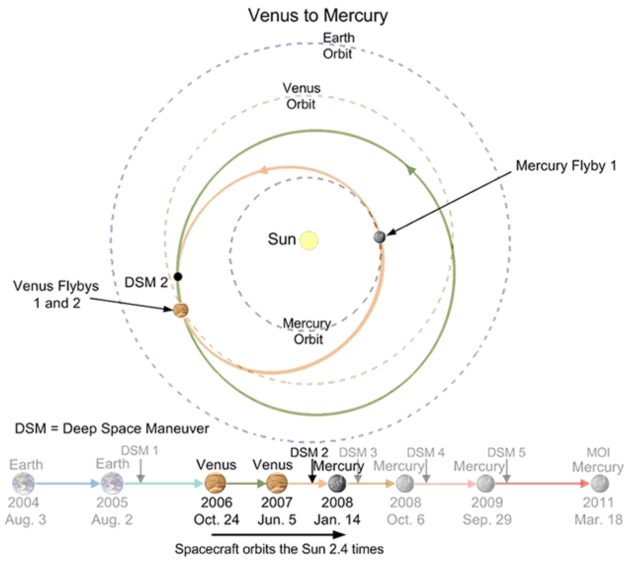


Figure 1: Trajectory of the Mercury probe. [Figure courtesy of NASA/JHU Applied Physics Laboratory/Carnegie Institution of Washington [7]]

The part of the MESSENGER orbit before the flyby (marked by the arrow on the right hand side) is shown in pale red – it's a nice elliptical orbit. Hence, it seems reasonable to apply the relativistic perihelion shift formula to both the MESSENGER and the Mercury orbits, and to see what comes out.

The angular shift of the perihelion [8] per revolution as stemming from relativistic corrections to Newtonian gravitation is given by

$$\delta\phi = 6\pi \frac{GM}{c^2} \frac{1}{a(1-e^2)}. \quad (1)$$

Here a is the so-called semi-major axis [9] of the orbit (that's half the longer diameter of the ellipse), and e is its eccentricity (for a circle, $e = 0$); the larger e is, the more elongated the ellipse. Sometimes, the quantity $a(1-e^2)$ is called the semi-latus rectum, L [10] (sorry, the geometry of conic sections is pretty old, hence all the Latin and Greek). Since the perihelion distance, p, is related to the semi-major axis by $p = a(1-e)$, we can also write

$$\delta\phi = 6\pi \frac{GM}{c^2} \frac{1}{p(1+e)}. \quad (2)$$

The fraction GM/c^2 is half the so-called Schwarzschild radius [11] – for the Sun, $GM/c^2 \approx 1.5\,\mathrm{km}$. Since this is very small compared to Mercury's perihelion distance of 46 million km, the perihelion shift per revolution is a tiny angle.

However, what we actually want to know when we are interested in navigating a spaceprobe very close to Mercury is not this angle, but the specific motion of the perihelion measured in kilometres. This motion then quantifies the offset Δ along the orbit due to relativistic effects. But it is easy to obtain: Since the angle is given in radians, we just have to multiply by the perihelion distance, p, and obtain

$$\Delta = p\delta\phi = 6\pi \frac{GM}{c^2} \frac{1}{1+e} \approx \frac{30\,\mathrm{km}}{1+e}. \quad (3)$$

Curiously, this offset depends only on the eccentricity, and is there even if the orbit is a perfect circle!

Now, we can apply this formula to the orbits of Mercury and MESSENGER and plug in some numbers: The eccentricity of Mercury is $e = 0.20$, which yields Δ = 25 km. This is the relativistic offset of the orbit that accumulates over the 88 days of one revolution. Now, however, we are not dealing with an entire revolution. But the crucial point is that for shorter periods, we can just take the respective part of this shift. Thus, for the 26 days between the trajectory correction manoeuvre 19 (TCM-19) on December 19 and the MESSENGER flyby on January 14, the estimate for the relativistic offset amounts to about 7 km.

The orbital parameters of MESSENGER can be obtained from the Jet Propulsion Laboratory's (JPL) Horizons web site [12]. This is a very cool interactive site where you can get all kinds of coordinates for nearly all the Solar System! The elliptic orbit MESSENGER was on around the Sun on, say, January 1, 2008 had an eccentricity $e = 0.38$, and a period of about 140 days. The relativistic offset Δ of this orbit amounts to 22 km, and scales down, for 26 days, to about 4 km.

At this point, I should add that the estimate using the perihelion precession formula does not take into account the timing along the trajectory, i.e. the actual moment when the perihelion is reached. This timing is different in the Newtonian and General relativistic theory, because of the time dilatation (or, gravitational redshift) in the gravitational field of the Sun.

Nevertheless, taken all together, the relativistic effects on the MESSENGER trajectory over the period of 26 days since the last correction manoeuvre can add up to an uncertainty of a few kilometres at the moment of the Mercury flyby – an offset that matters.

Indeed, upon closer inspection I realised that General Relativity comes into play when manoeuvring an interplanetary spacecraft at different points these days. General Relativity

- enters the very definition of the barycentric coordinate system used to describe motion in the Solar System [13],
- is taken into account to calculate the ephemerides [14] (the coordinates as a function of time) of the planets and other Solar System bodies, and
- is taken into account to calculate trajectories of space probes [15].

Regarding the last point, a lot of theoretical background and useful information can be found in "Monograph 2: Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation" by Theodore D. Moyer [15], freely available at the web site of the JPL. It shows, for example, a very impressive set of relativistic equations of motion (page 4-19 of [16]) that have been implemented in the JPL software to calculate trajectories.

In the meantime, I had received long and informative answers to my many questions in an email from Tony Taylor. Tony is a veteran navigator who has steered space probes to all planets of the Solar System besides Pluto. He now works for kinetX [17], a company founded by former JPL engineers like himself that offers consulting and service to commercial satellite programs, but also has contracts with NASA to do trajectory calculations for several interplanetary missions, including MESSENGER.

Tony told me that the software they use can indeed account for all the general relativistic terms described by Moyer. However, these terms actually are switched on only for the Sun, and for Jupiter in case a spacecraft gets in the close vicinity of this planet. For the MESSENGER mission, General relativistic corrections are taken into account for the Sun only. A calculation of the MESSENGER trajectory with these corrections to the Newtonian gravitational field switched on for all planets yields the same results, within the margin of numerical errors.

Including the general relativistic terms for the central potential of the Sun, however, is important. Tony was so generous and took the time to play around a bit with his code to quantify the relativistic effects in the calculation of the MESSENGER trajectory.

He integrated the spaceprobe's trajectory without relativistic corrections, starting from a point at September 30th up to the Mercury flyby in January. The resulting flyby altitude at Mercury increased by about 10 km as compared to the calculation including the relativistic correction for the gravitational potential of the Sun, and the time of the closest approach changed by about 13 seconds, which corresponds to a distance along the path of MESSENGER of about 65 km. Neglecting General Relativity, argues Tony, may not have been disastrous to the mission, but it would have caused some consternation in mission control about what was going on.

A second comparison was done for trajectories starting December 19th – the day of the correction manoeuvre TCM-19 – calculated with and without relativistic corrections, yielding a difference considerably smaller, amounting to about 250 m in altitude and about 5 km along the trajectory. Note that this is less than the estimate based on the perihelion shift formula, a difference mainly due to the different timing of the passage along the orbit.

Finally, Tony compared the post-flyby heliocentric orbit of MESSENGER – which has an eccentricity of 0.3816 – as calculated with and without relativistic corrections to the Sun's gravitational field. He found that the spacecraft orbit shows a perihelion shift due to relativity of about 24 arcsec per century, in good agreement with the theoretically expected value.

The extra acceleration of MESSENGER due to General Relativity describing the gravitational field of the Sun is tiny, but not negligible. At the aphelion of the current post-flyby spacecraft orbit, it amounts to about $0.5\times^{-6}$ mm/s^2, and at perihelion, it's roughly ten times larger.

However, to put this in perspective, Tony told me that the acceleration due to the pressure of solar radiation is about $150\times^{-6}$ mm/s^2 at aphelion, and 450×10^{-6} mm/s^2 at perihelion. Thus, the relativistic effect – which does not follow the $1/r^2$ law – varies only between 0.4 to 1.2 percent of the solar pressure effect! However, the velocity tracking of the spaceprobe by Doppler data is sensitive to about 0.1 mm/s, meaning that neglected relativistic corrections would show up in the data within a few days.

Curiously enough, with all this precision now available, the MESSENGER team before the flyby was a bit worried about Mercury: In all the spacecraft trajectory calculations mentioned so far, the positions and velocities of the planets – the ephemerides, as they are called – have to been taken for granted. The MESSENGER team relies on the ephemerides DE-405 as supplied by JPL, which again are established on the basis of astronomical observations and spacecraft flyby data. Now, Mercury has been visited only once, by Mariner 10 about thirty years ago, and the last astronomical observation used for DE-405 is more than ten years old. So, essentially the MESSENGER navigators had to trust these ephemerides and hope that Mercury would be found where expected: Due to the small distance to Mercury in the planned flyby, any unaccounted-for large systematic error in the planets ephemerides could have resulted in a "landing" instead of a flyby. There was an emergency manoeuvre planned at 34 hours before the flyby to cover this contingency. Fortunately, as it came out, this manoeuvre was not necessary. In fact, the MESSENGER flyby data are now being used to check and improve the Mercury ephemerides.

In the end, I've learned that along general lines, Newton is still safe enough a pilot for MESSENGER. However, when it comes to the fine-tuning of the trajectory, Einstein's correcting hand is clearly visible.

Endnotes

☿ Calculating orbits around a central mass in General Relativity amounts to calculating geodesics in the Schwarzschild metric [18]. It turns out that, for large enough distances from the centre, the motion of a mass corresponds very well to the motion according to Newtonian gravity, but in addition to the $1/r$ potential of Newton, there is an extra term, proportional to $1/r^3$. This extra term acts as a perturbation to the Newtonian elliptic orbits, and yields a shift of the perihelion of the ellipse.

☿ The application of General Relativity to the calculation of trajectories of spacecraft is realised within a theoretical framework known as the Parametrized Post-Newtonian (PPN) formalism [19] of Will and Nordtvedt (1972) [20]. The full set of equations used to describe the acceleration of a mass i due to the gravitational influence of other masses j then reads (page 4-19 of "Monograph 2: Formulation for Observed and Computed Values of Deep Space Network Data Types for Navigation" by Theodore D. Moyer [16])

$$\begin{aligned}\ddot{\mathbf{r}}_i = \sum_{j\neq i}\frac{\mu_j(\mathbf{r}_j-\mathbf{r}_i)}{r_{ij}^3}\Bigg\{ & 1-\frac{2(\beta+\gamma)}{c^2}\sum_{l\neq i}\frac{\mu_l}{r_{il}}-\frac{2\beta-1}{c^2}\sum_{k\neq j}\frac{\mu_k}{r_{jk}} \\ & +\gamma\left(\frac{\dot{s}_i}{c}\right)^2+(1+\gamma)\left(\frac{\dot{s}_j}{c}\right)^2-\frac{2(1+\gamma)}{c^2}\dot{\mathbf{r}}_i\cdot\dot{\mathbf{r}}_j \\ & -\frac{3}{2c^2}\left[\frac{(\mathbf{r}_i-\mathbf{r}_j)\cdot\dot{\mathbf{r}}_j}{r_{ij}}\right]^2+\frac{1}{2c^2}(\mathbf{r}_j-\mathbf{r}_i)\cdot\ddot{\mathbf{r}}_j\Bigg\} \\ & +\frac{1}{c^2}\sum_{j\neq i}\frac{\mu_j}{r_{ij}^3}\left\{[\mathbf{r}_i-\mathbf{r}_j]\cdot[(2+2\gamma)\dot{\mathbf{r}}_i-(1+2\gamma)\dot{\mathbf{r}}_j]\right\}(\dot{\mathbf{r}}_i-\dot{\mathbf{r}}_j) \\ & +\frac{3+4\gamma}{2c^2}\sum_{j\neq i}\frac{\mu_j\ddot{\mathbf{r}}_j}{r_{ij}}\end{aligned}$$

The first term is the Newtonian term, and the first term in the curly brackets is the first correction by General Relativity, which yields the perihelion shift. All other terms are much smaller (some of them depend on the velocity – that's what is called gravitomagnetism [21] and yields the Thirring-Lense effect [22], among others). Actually, the PPN formalism allows to describe the effects of many different hypothetical gravitational theories by introducing the parameters β and γ into the equations. Newtonian gravity corresponds to the choice $\beta = \gamma = 0$ and General Relativity to $\beta = \gamma = 1$. Comparing calculated trajectories with actual data can then even be used as a test of General Relativity. This is done by varying β and γ in such a way that the calculation fits best to the data. So far, General Relativity provides the best fit, and for practical applications, β and γ are set to their general relativistic values.

☿ It is important not to forget that there are other forces besides gravitation acting on spaceprobes. The influence of light pressure is about one hundred times larger than the corrections to gravitation due to General Relativity, and one might wonder about the role of the Solar wind. However, despite its dramatic name, the Solar wind can safely be neglected: Solar wind, which is now monitored by satellites such as the NASA ACE probe [23], has an effect about 5000 times smaller than light pressure. Thus, it is negligible even compared to the tiny corrections by General Relativity.

References

[1] http://www.nmspacemuseum.org/halloffame/detail.php?id=71

[2] http://messenger.jhuapl.edu/index.php

[3] http://messenger.jhuapl.edu/gallery/sciencePhotos

[4] http://www.planetary.org/blog/article/00001300

[5] http://messenger.jhuapl.edu/the_mission/maneuvers.html

[6] http://en.wikipedia.org/wiki/Tests_of_general_relativity#Perihelion_precession_of_Mercury

[7] http://messenger.jhuapl.edu/the_mission/trajectory.html

[8] http://en.wikipedia.org/wiki/Kepler_problem_in_general_relativity#Precession_of_elliptical_orbits

[9] http://en.wikipedia.org/wiki/Semi-major_axis

[10] http://en.wikipedia.org/wiki/Semi-latus_rectum

[11] http://en.wikipedia.org/wiki/Schwarzschild_radius

[12] http://ssd.jpl.nasa.gov/horizons.cgi

[13] http://arxiv.org/abs/astro-ph/0303376

[14] http://adsabs.harvard.edu/abs/1996DPS....28.2504G

[15] http://descanso.jpl.nasa.gov/Monograph/series2_section.cfm

[16] http://descanso.jpl.nasa.gov/Monograph/series2/Descanso2_S04.pdf

[17] http://www.kinetx.com/

[18] http://en.wikipedia.org/wiki/Schwarzschild_metric

[19] http://en.wikipedia.org/wiki/Parameterized_post-Newtonian_formalism

[20] http://adsabs.harvard.edu/cgi-bin/bib_query?1972ApJ...177..757W

[21] http://en.wikipedia.org/wiki/Gravitomagnetism

[22] http://en.wikipedia.org/wiki/Frame-dragging

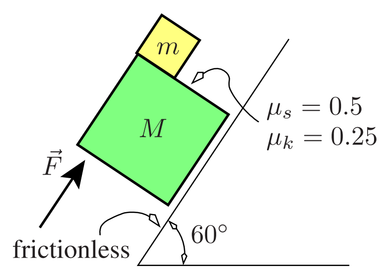
[23] http://www.srl.caltech.edu/ACE/

Stefan Scherer is an editor for the Landolt-Börnstein Compilation of Data and Functional Relationships form physics and related sciences. He can be reached at scherer@th.physik.uni-frankfurt.de. This article is adapted from two posts that originally appeared on Stefan and Sabine's blog, http://backreaction.blogspot.com/.

The SIN Bin

Rohan Jayasundera

SIN BIN #124



Two masses, m and M, are pushed upward along a frictionless inclined surface with a force F as shown in the figure. Calculate the minimum force, F, that prevents m from sliding downward on M. [$\mu_s = 0.5$, $\mu_k = 0.25$, $m = 1$ kg, $M = 9$ kg] Let $g = 10$ m/sec^2. Answer in Newtons.

SIN BIN #123 – Solution

An object, mass m is connected to a rope (of negligible mass) and rotates in a vertical plane of radius $R = 2m$. At the position shown in the graph, the magnitude of the total acceleration (a), the tension force (T), and the speed (v) are given as 7 m/sec^2, 26 N and 3 m/sec respectively. Find the mass of the object. Answer in kg.

The total acceleration is the vector sum of the radial acceleration and the tangential acceleration.

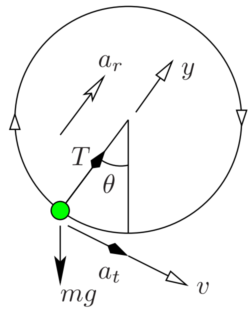
$$\vec{a} = \vec{a}_r + \vec{a}_T$$

$$|\vec{a}| = \sqrt{|a_r|^2 + |a_T|^2} = \sqrt{49} = 7 \text{ m/s}^2$$

Since the speed at the instant shown is, $v = 3$ m/s

$$|a_r| = \frac{v^2}{r} = 9/2$$

If we use Newton's second law in the positive "r" direction as shown



$$\sum F_x = ma_x$$

$$mg\sin\theta = ma_t$$

$$a_t = g\sin\theta \qquad (1)$$

If we do the same in the "y" direction

$$\sum F_y = ma_y = ma_r$$

$$T - mg\cos\theta = ma_r = \frac{mv^2}{r}$$

$$T - mg\cos\theta = m\frac{9}{2} \qquad (2)$$

Since $|a|^2 = |a_r|^2 + |a_t|^2$

$$49 = (9/2)^2 + (g\sin\theta)^2$$

$$\frac{49 - (9/2)^2}{(9.8)^2} = \sin 2\theta$$

$$\therefore \sin\theta = \frac{\sqrt{49 - (9/2)^2}}{9.8}$$

$$\therefore \theta = 33.17^\circ$$

Now we could substitute this value into equation (2)

$$T - mg\cos\theta = m(9/2)$$

$$\therefore m = \frac{26}{g\cos\theta + 4.5} = 2.0 \text{ kg}$$

Sudoku Puzzle

Anthony Anderson

Here's a twist on the addictive game: letters instead of numbers! Fill every row, column and 3×3 box, so that each contains the nine different letters forming the surname of a Nobel prize winning physicist. When you have completed the puzzle, use the numbered letters in the grid to complete the surname and Nobel citation of this scientist.

[9][6][2][4][3][14][11][1][5]

F[1]R [4][6][11] [12]PT[8][2]A[14] PR[3][10][13][16][34][15][5]
[47][17][19]TRU[9][18][22]T[19] A[24]D T[7][26]
[27]P[33][29]TR[15][32][38][20]P[6][41] A[31]D
[25][42]TR[20][23][28]G[8][44]A[37] [13][5]V[46][11]T[34]GAT[47][39][17][16]
[2]ARR[6][3]D [1]UT W[8]T[21] T[35][18][13]R A[34]D.

The solution to this puzzle will appear in the next issue of Phys 13 News.

S	1	2	3	4	5	M	6	L
7	E	N	8	9	L	10	11	12
13	M	14	15	16	C	H	E	17
18	19	20	L	I	21	22	M	C
23	24	25	C	26	27	I	28	H
29	H	I	30	31	O	32	L	33
N	34	E	35	O	36	37	38	S
39	40	S	N	41	I	42	43	M
M	44	45	S	L	46	O	N	47

Anthony Anderson is a Distinguished Professor Emeritus in the Department of Physics & Astronomy at the University of Waterloo. He can be reached at aanderson@scimail.uwaterloo.ca.

Prof Quotes

George McBirnie

A compilation of recent memorable quotations taken from actual physics lectures at the University of Waterloo.

Prof: "We have Alice here (A on top of first stick person) and her twin sister Alice prime (A') over here."
–Epp, Phys 490 (Introduction to Gravitational Physics)

Prof: "On average, Betelgeuse doesn't exist."
–Balogh, Phys 375 (Astrophysics 2 – Stellar Evolution)

Prof: "If you don't do the assignments, you won't do well on the exam. You will get a worse mark and then you will have to work at Burger King."
–Forrest, Phys 263 (Classical Mechanics & Special Relativity)

Prof: "Once nano becomes not sexy, you can still use physics."
Student: "Why are we building a nano building then?"
Prof: "Because it uses very little material."
–Forrest, Phys 263 (Classical Mechanics & Special Relativity)

George is a fourth year undergraduate student in the Department of Physics & Astronomy at the University of Waterloo. He has been collecting prof-quotes for years and plans on eventually publishing a book. He can be reached at gmcbirni@sciborg.uwaterloo.ca.

Subscription Form for Phys13News

Name: ....

Street: ....

City .... Province/State: ....

Country: .... Postal/Zip Code ....

Rates: four issues per year

	Annual Subscription	3 Year Subscription	
Canada	$12 Canadian	$30 Canadian	7% GST included
USA	$15 US	$35 US	(R119260685)
Other	$18 US	$40 US	

A limited number of reprints are available on request. A larger number of reprints (minimum of 25) can be provided at an additional cost. Please inquire by email.

Make cheque or money order payable to **Phys 13 News** Amount Enclosed ....

Please send your completed order form and await your next issue. If you really need a receipt or an invoice, add a service charge of $1 to the amount of your subscription and check here: ......

Is this a new subscription or a renewal order? ......

Return undeliverable Canadian addresses to:

Phys13News
University of Waterloo
Department of Physics & Astronomy
200 University Avenue West
Waterloo, Ontario N2L 3G1

Return postage guaranteed
University of Waterloo Publication Number 40065122