The attractions of magnetism for nanoscale data storage

R. P. Cowburn

Phil. Trans. R. Soc. Lond. A 2000 **358**, 281-301
doi: 10.1098/rsta.2000.0532

The attractions of magnetism for nanoscale data storage

By R. P. Cowburn

Nanoscale Science Group, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK (rpc12@cus.cam.ac.uk)

The latter half of the 20th century was marked by a ten-million-fold increase in the engineer's ability to store information magnetically (e.g. computer hard-disk drives). Without this, the computing revolution that the developed world has experienced would not have happened. In this paper, I discuss the technical difficulties that face engineering if it is to continue storing increasingly large amounts of information into the third millennium and show how applications of the emerging fields of nanotechnology and quantum engineering may provide solutions. I offer predictions for the course of development of information-storage technology, including concepts such as the quantized magnetic disk, magnetic RAM chips and new magnetic materials made from artificial giant atoms. The changes that might come about in society as a result of mass information storage are discussed.

Keywords: magnetism; memory; nanotechnology; quantum mechanics; data storage; hard-disk drive

1. Introduction

Magnetism has been known to mankind for many hundreds of years and has, from the beginning, found an important use in navigation. In the first century BC, scholar Lucretius wrote of the magnetic properties of Lodestone, although his understanding of magnetism was perhaps better informed than that of the 13th century AD scholar Bartholomew the Englishman, who assures us that 'This kind of stone restores husbands to wives and increases elegance and charm in speech. Moreover, along with honey, it cures dropsy, spleen, fox mange, and burn' (Verschuur 1993).

The London physician William Gilbert was the first to make serious inroads into an understanding of magnetism. A Fellow of St John's College, Cambridge, Physician to Elizabeth I and the founder of a precursor to the Royal Society, Gilbert published his great work '*De Magnete Magneticisque Corporibus et de Magno Magnete Tellure Physiologia Nova*' ('On the magnet: magnetic bodies also, and on the great magnet the Earth; a new physiology') in 1600. Gilbert's eminent contemporary Galileo Galilei read the work and consequently described Gilbert as being 'great to a degree that is enviable'. The portrait shown in figure 1*a* was left to Oxford University after his death and bears the inscription 'Gilbert, the first investigator of the powers of the magnet'.

Gilbert lived too soon, however, to see the real story of magnetism unveiled; this had to wait another 200 years for the arrival of Maxwell and Faraday. Their great contribution was to unite electricity and magnetism into 'electromagnetism'. Another

Phil. Trans. R. Soc. Lond. A (2000) **358**, 281–301

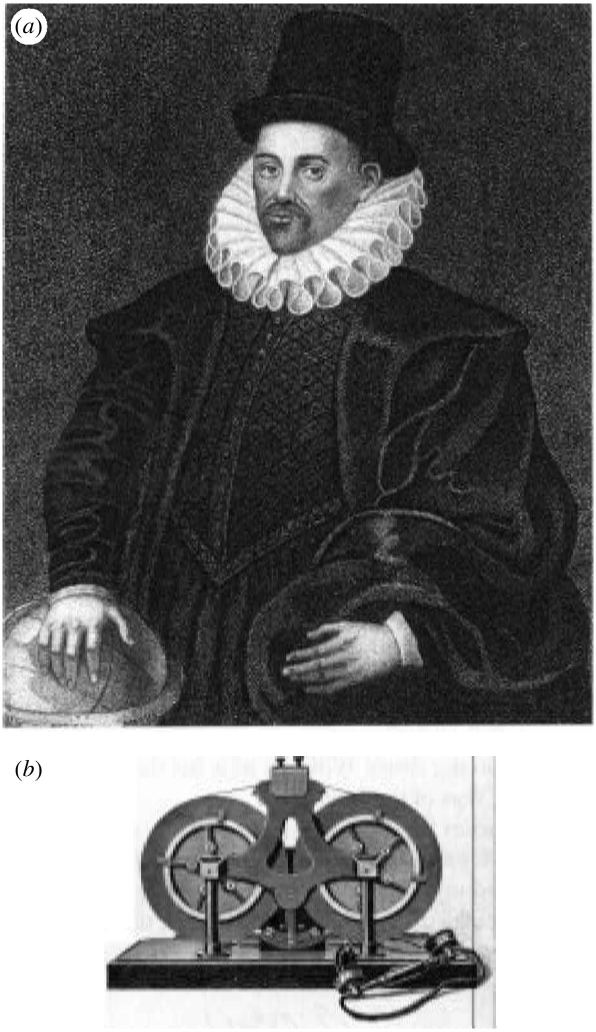

Figure 1. Views from antiquity. (*a*) William Gilbert, the father of magnetic research; (*b*) Poulsen's telegraphone (from *Scientific American*, 22 September 1900).

revolutionary application thus appeared from magnetism: the dynamo, or electric generator, and its sister machine, the electric motor. It is from these that the 20th-century remodelling of society by electricity and electronics flowed.

In addition to navigation and electromagnetism, magnetism has impacted human history in a third way. This third revolution was started very quietly at the end of the 19th century in the laboratories of the Copenhagen Telephone Company by Valdemar Poulsen. Poulsen reasoned that people would find a device that could record telephone messages useful and so invented the 'telegraphone' (figure 1*b*), a precursor to the tape recorder but using piano wire instead of magnetic tape. The Austrian Emperor Franz Joseph saw the telegraphone at the Paris Exposition in 1900 (where, incidentally, it won the Grand Prix) and recorded a message. That message still exists today, and is the world's oldest magnetic recording.

The irony is that the telephone recording application envisaged by Poulsen failed to materialize in most people's homes until this decade, some 100 years later. Nevertheless, the principle that magnetism could be used as a *memory* device had been demonstrated. This principle is now used in every aspect of modern life, be it storing information on your computer, storing bank records, recording hospital scans, preparing television and radio programmes, preparing music CDs, or saving personal details on the backs of credit and debit cards.

2. Magnetic memory

Scientists express the important concept that magnetic materials possess a memory function by a *hysteresis loop*. Figure 2*a* shows a hysteresis loop measured in our laboratory from a very common magnetic material called permalloy ($Ni_{80}Fe_{20}$). A hysteresis loop is a graph of *magnetization* (a measure of the extent and direction that a material has been magnetized) against a *magnetic field*. The most important feature of this loop is that under zero applied magnetic field, the magnet can have two different magnetization states (called the remanent states): fully positive or fully negative. Which state is obtained depends upon the field *history*. The fact that the remanent state of the magnet depends upon the fields to which it has been exposed in the past is precisely the principle used by all magnetic recording, be it audio cassettes, video tape, computer floppy disks or computer hard disks.

The first question anybody asks of a new magnetic storage technique is 'what is its storage density?' This means how many bits of information (1s and 0s) can be stored in a given area, typically a square inch. This is important, because the maximum size of the disk or cassette is usually limited and so the storage density determines how much information can be stored. The most striking feature of the magnetic storage industry is the extent to which the storage density has increased during the last 40 years. Figure 2*b* shows hard-disk storage density plotted against time between 1956 and 2000 (after Grochowski & Thompson (1994)). Within this time window, storage densities have increased by a factor of ten million. The growth is mathematically exponential, with a current annual compound growth rate of 60%. To put this in perspective, if in 1956 our ability to store data is likened to writing a large, single letter on a sheet of A4 paper, such as a young child might do, then today we have learned to write letters 0.1 mm high, which is fine enough to copy the entire Holy Bible onto a single sheet of A4 paper. In 18 months' time we shall be able to do it on half a sheet of A4 and so on. Such improvement in performance without changing paradigm (today's best hard disks use the same principle of operation as Poulsen's telegraphone) is, to the best of my knowledge, unparalleled in any other sphere of engineering, except perhaps by the development of the microchip. Together, these are undoubtedly two of the great technical achievements of the 20th century. The question that ends the second millennium and which will dominate the early part of the third millennium is this: how to keep the growth going?

Figure 3 shows some modern data storage disks and the inside of a state of the art hard-disk drive. This image covers approximately a 10 year span and shows an increase in capacity from 100 kb to 19 Gb; 1 Gb (gigabyte) is 1000 Mb (megabytes), 1 000 000 kb (kilobytes) and 1 000 000 000 bytes; a byte is an integer number between 0 and 255. The hard drive is based around a large circular disk that is ungrammatically called the 'media', and consists of a rigid plate on which a thin film of a

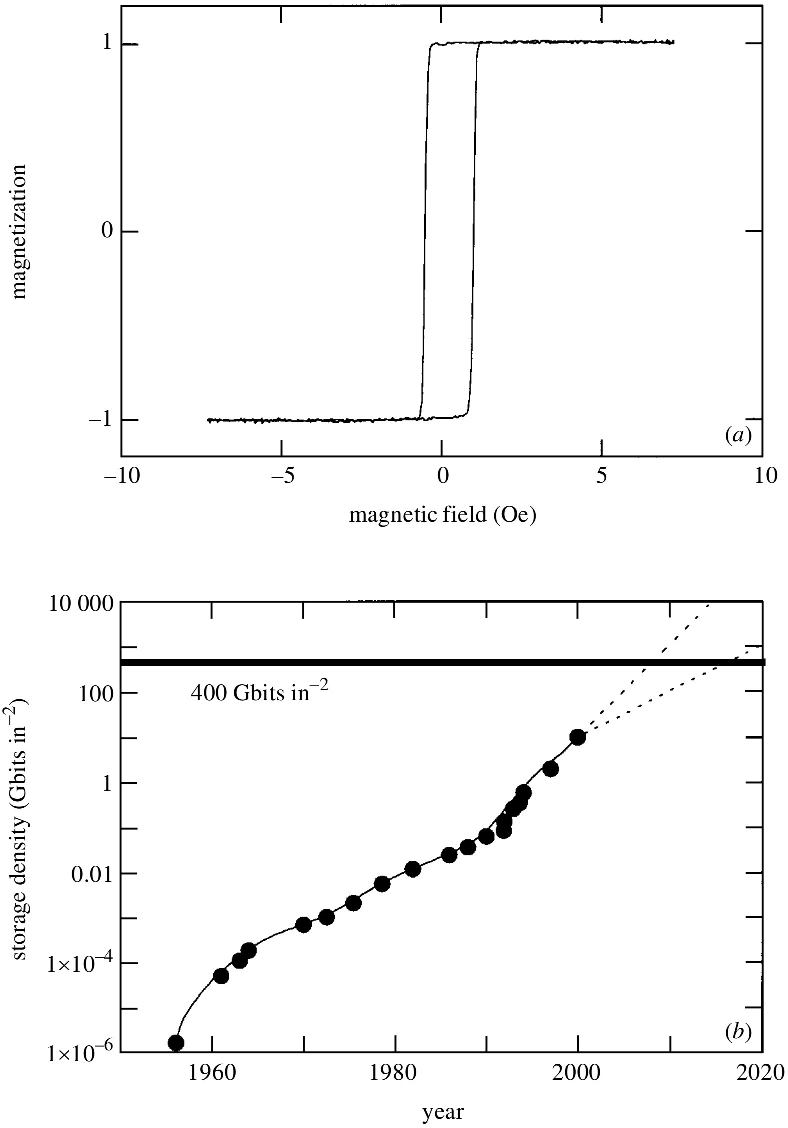

Figure 2. The principle and practice of magnetic storage. (*a*) A hysteresis loop for permalloy; (*b*) hard-disk storage density during the last 50 years.

cobalt–chromium alloy has been deposited by a technique called 'sputtering'. The disk spins at *ca.* 5000 rpm. In the foreground is an arm with a small sensor at the end, called the 'head'. A combination of the disk spinning and the arm rotating allows the head to reach any part of the disk surface. The head comprises two parts: a write head, used to magnetize the disk locally, either positively or negatively, and a read head, used to sense the magnetic field coming from the disk and hence determine the direction (fully positive or fully negative) in which a region has been previously magnetized. These two parts, respectively, achieve the data recording and playback functions.

Figure 3. A survey of modern storage technology. From the left, an old 5.25″ floppy disk (100 kb), a modern 3.5″ disk (1.4 Mb), a Zip disk (100 Mb), a magneto-optical disk (up to 640 Mb). In the centre is an open hard-disk drive (up to 27 Gb).

A stream of 1s and 0s is thus written to the disk as a circular track of positive and negative magnetization in a row. Figure 4 shows an actual picture taken by a special microscope (called a magnetic force microscope) that can 'see' positive and negative magnetization. Each magnetized region corresponds to one data bit, and so the storage density of the disk is determined by how large each of the written bits is. At the time of writing, the most advanced hard disks are writing bits *ca.* 100 nm in size (1 nm is a thousand millionth of a metre, or approximately the length of 10 atoms). The phenomenal growth in storage density described previously has, up until now, been achieved simply by reducing the dimensions of the written bit by 70% in both width and length every 18 months, to give an overall doubling of density. However, fundamental physical limitations, which are expected to prevent the bit size shrinking much further, now stop engineers from sleeping as soundly as they used to.

The first of these limitations is due to the material from which the media is made and is known as the 'media noise problem'. Hard-disk media is not made from a uniform sheet of magnetic material. Rather, it comprises a large number of small magnetic islands (or grains) tightly packed together in a non-magnetic sea. The small non-magnetic regions between the grains prevent one bit from 'leaking' and growing in size. If the bit size is large compared with the grain size, then a well-defined rectangular bit is possible. If, however, the bit size is comparable with the grain size, then the bit has rough, ill-defined edges, which increase the noise in the sensor. The number of erroneously read bits therefore increases.

One approach to the problem of media noise is to find new materials for the media that have smaller grains. This, however, leads on to a second problem, called 'superparamagnetism', which comes about because of temperature. Just as a molecule at a finite temperature vibrates and rotates, so the magnetization direction in a magnetic material must fluctuate. Now, in a large bar magnet this does not matter because the size of the fluctuation is tiny. The smaller the magnet becomes, however, the larger

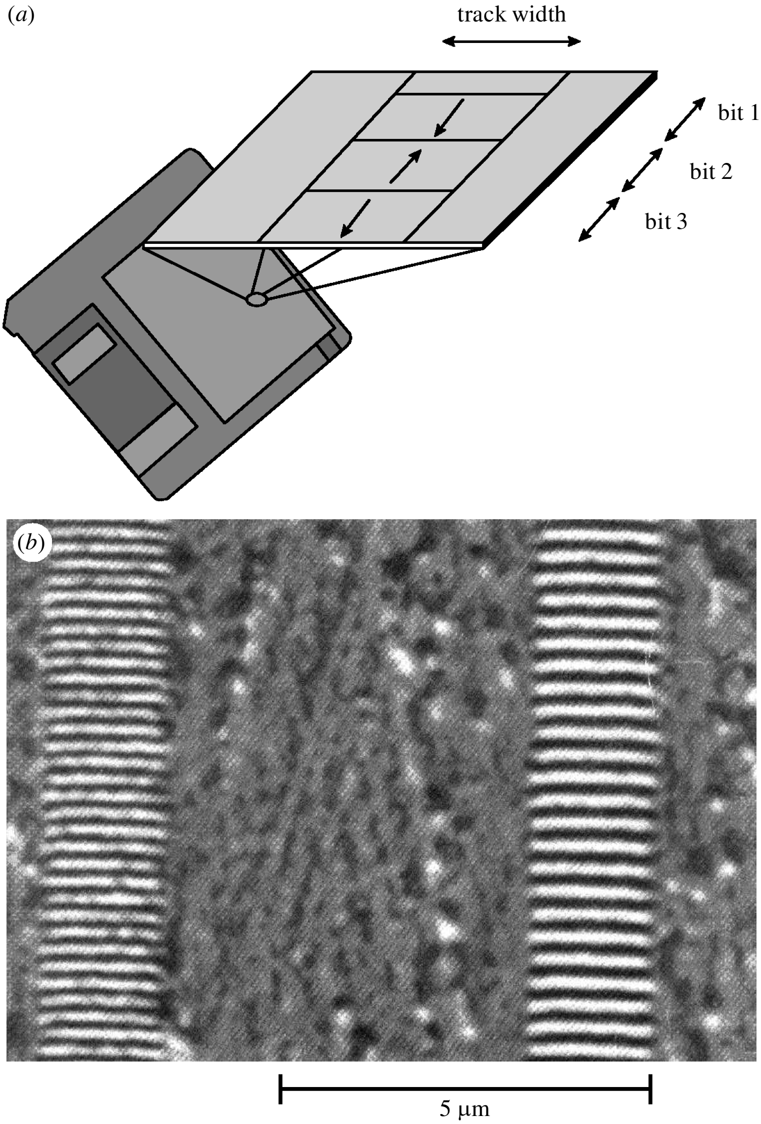

Figure 4. Magnetic disk storage. (*a*) A schematic representation of written bits; (*b*) a magnetic force microscope image of two tracks of written bits (from Malhotra *et al*. (1997), © 1997 IEEE).

the fluctuations are in absolute terms. Current hard-disk media grains are still large enough that the magnetization fluctuations are not too important (although one of the consequences of thermal fluctuations, even in today's technology, is that floppy disks will lose their memory after a year or so, a fact about which the manufacturers are strangely silent). If, however, the grain size is reduced much further, then the magnetization fluctuations become so large that they can sometimes turn the magnetization through 180°, converting a 1 to a 0 and corrupting the data. Superparamagnetism is expected to become a serious problem by the year 2005 if grain sizes continue to shrink at the current rate.

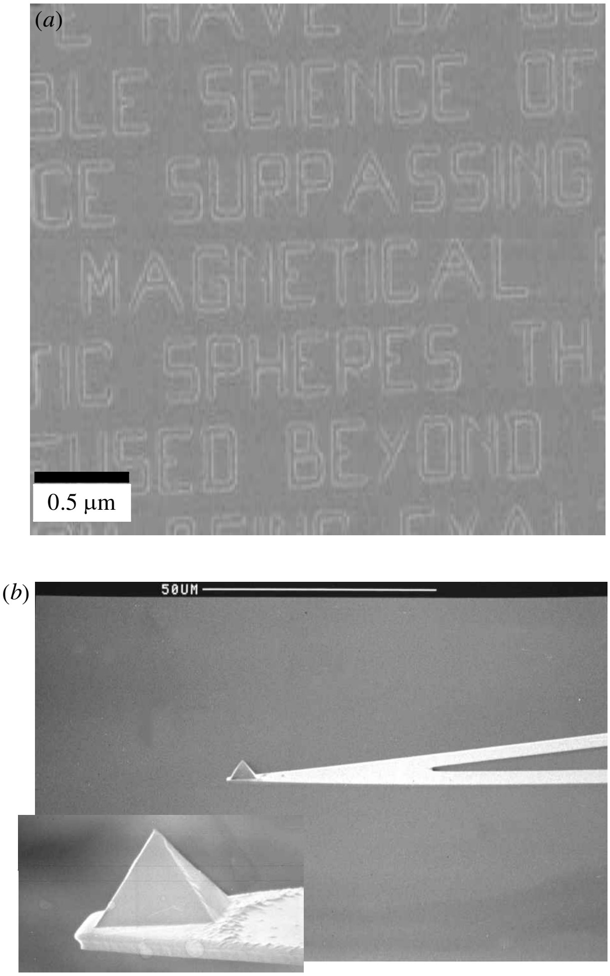

Figure 5. Excerpts from nanotechnology. (*a*) Part of Gilbert's *De Magnete* written in our laboratory by electron-beam lithography in letters 250 nm high using penstrokes 30 nm wide; (*b*) a scanned probe cantilever with a close up of the sharp tip inset (picture by J. Barnes).

A third problem is concerned with reading the data back from the disk. The smaller the bit size, the smaller the magnetic field that comes out of it. An increasingly sensitive detector is, therefore, needed for the read head, and the read head must 'fly' increasingly close to the surface of the disk, which causes tribological difficulties.

3. Magnetic nanotechnology

Nanotechnology is a field with immense future potential that has arisen largely during the 1990s. It is the art and science of manipulating and using material on the nanometre scale. There are two main workhorses in nanotechnology: the electron microscope and the scanned probe. The electron microscope is used to perform electron-beam lithography, which is essentially a photographic process. Instead of producing glossy pictures, however, it produces physical structures by using the sub-nanometre size focused electron beam to draw out the shape of the desired nanostructure onto a photographic layer. A number of chemical steps, directly analogous to developing a camera film, are then used to convert the exposed shapes into real physical structures, which, for the current state of the art, can be as small as 5 nm. Figure 5*a* shows an example of electron-beam lithography, where a passage from Gilbert's *De Magnete* has been written onto a piece of silicon. Each letter is 250 nm high and is written with penstrokes of width 30 nm. This is small enough to copy all 20 volumes of the Oxford English Dictionary into the space occupied by a single capital letter on this page!

The scanned probe is a development of the scanning tunnelling microscope (STM†), which won Binnig and Rohrer half of the Nobel Prize for Physics in 1986 for work that they carried out in the IBM Zurich research laboratory. In STM, and its sister technique of atomic force microscopy (AFM), an ultra-sharp point (see figure 5*b*) is used to 'feel' the shape of an object as it is scanned over a surface. Single atoms can be seen with these techniques, and so scanned probe microscopy is an excellent way of seeing nanostructures. Moreover, the sharp point can be used to pick up and move pieces of material as small as a single atom, and so STM can also be used for nanoscale building.

The technology of nanometre-scale construction is still not fully mature. One of the greatest challenges facing the field is that of construction speed. The text shown in figure 5*a* took *ca.* 10 s to write. While this may seem reasonably fast, some of the best commercial applications would need production times to be 100 million times faster than this. High-speed nanolithography will, therefore, need to be an important research topic during the third millennium if these devices are to leave the laboratory.

The reason for discussing nanotechnology in this article is because of a very important theoretical prediction made by William Fuller Brown Jr in 1968. Brown was one of the great pioneers of the field of *micromagnetics*, the mathematics of magnetism on a microscopic scale. In order to understand his prediction, we need to understand the different forces inside a magnet. In a typical magnet of any size there are usually at least two different internal processes, both competing with each other. These are called *demagnetization* and *exchange* and are illustrated in figure 6. In the case of demagnetization, the magnetization at the edges of a magnet 'sticks out' through the sides, leading to the appearance of magnetic north and south poles on the sides, just as in a conventional bar magnet. These poles are sources of magnetic field lines, which, as well as passing *outside* of the magnet (as seen by Faraday's famous iron-filing experiment), also pass backwards through the *inside* of the magnet. This field, called the *demagnetizing field*, opposes the magnetization direction and tries to rotate the magnetization in such a way as to reduce the strength of the poles. One way of doing this is by forming the whirlpool (or 'vortex') pattern shown in figure 6*a*. Such a pattern does not have any poles on the edges because the magnetization is always parallel to its nearest edge, and so the unfavourable demagnetizing field is not generated. If demagnetization were the only process at work in a magnet, then permanent magnets and magnetic memory would not exist, because the demagnetizing field could always cause the magnetization to collapse.

Fortunately, demagnetization is opposed by the second internal process, exchange. This is a quantum mechanical effect, and is ultimately responsible for magnetism.

---

† Also scanning tunnelling microscopy.

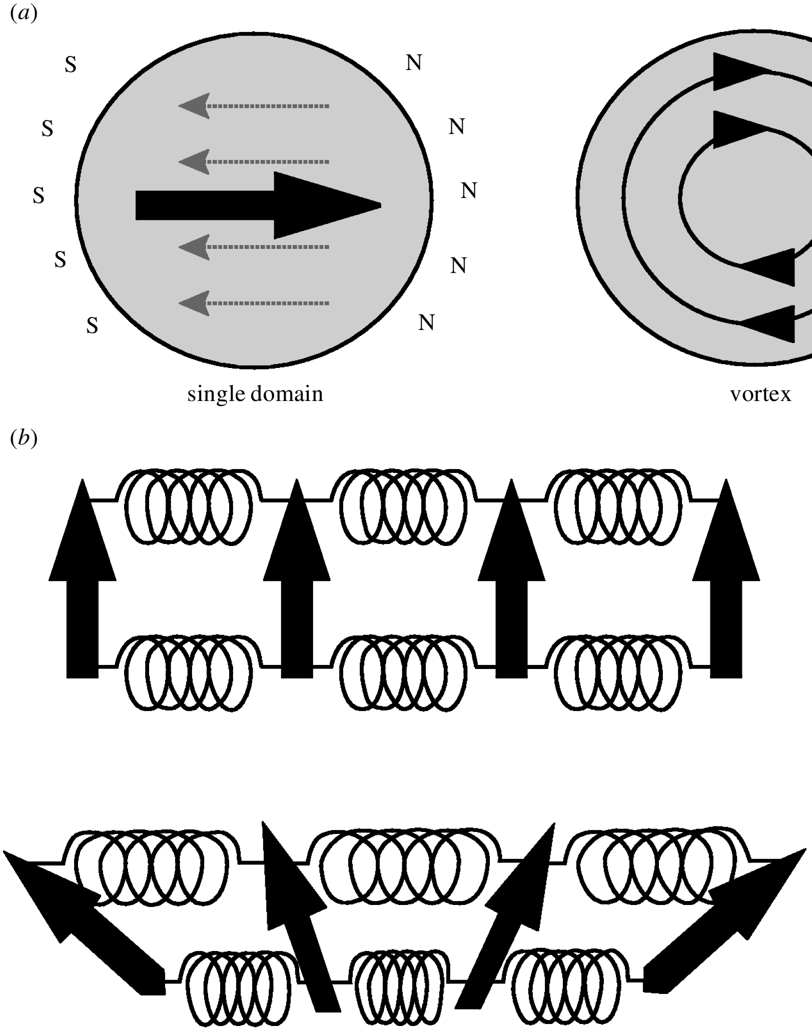

Figure 6. The two opposing forces in nanomagnets: (*a*) demagnetization, where the surface poles (N, S) create an internal demagnetizing field (dashed lines), which cause the magnetization (solid line) to break into a vortex; (*b*) quantum mechanical exchange, which appears as springs between spins, keeping neighbouring spins parallel.

Quantum mechanics is one of the great success stories of 20th-century physics, achieving popular fame via Schrödinger's cat, and is a highly counter-intuitive, but accurate, description of very small objects such as electrons and atoms. What we have so far referred to as 'magnetization' actually comes from a quantum property of electrons called 'spin'. Each electron can be thought of as a tiny bar magnet pointing either up or down, depending on the spin. One of the basic quantum laws is the 'Pauli exclusion principle', which says that no two objects can have the same quantum description as each other. This means that electrons of the same spin direction try to stay away from each other, which is actually a very good thing because, when two electrons do approach each other, there is an electrostatic energy penalty to pay. If,

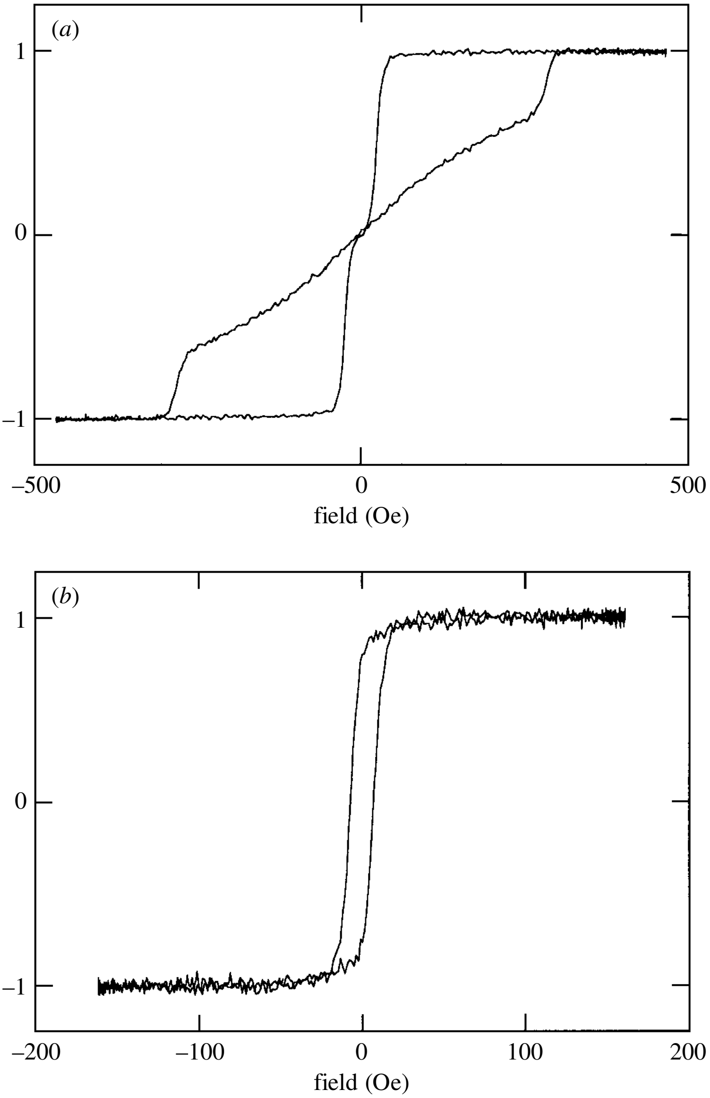

Figure 7. An experimental demonstration of Brown's fundamental theorem in circular permalloy nanomagnets. (*a*) A hysteresis loop from a 300 nm diameter nanomagnet showing the unfavourable vortex state; (*b*) a hysteresis loop from a 100 nm diameter nanomagnet showing the favourable single-domain state.

however, the electrons have their spins pointing in different directions, then the Pauli exclusion principle does not try to keep them apart because there is already some difference in their quantum descriptions. Consequently, the electrons encounter each other frequently and feel the full strength of the electrostatic repulsion that exists between two negative charges. Overall, the spins in a magnetic material attempt to stay aligned parallel with each other whenever possible in order to minimize energy. If one tries to push neighbouring spins to point in different directions, then it is as if there were springs between them (as shown in figure 6*b*). The exchange springiness,

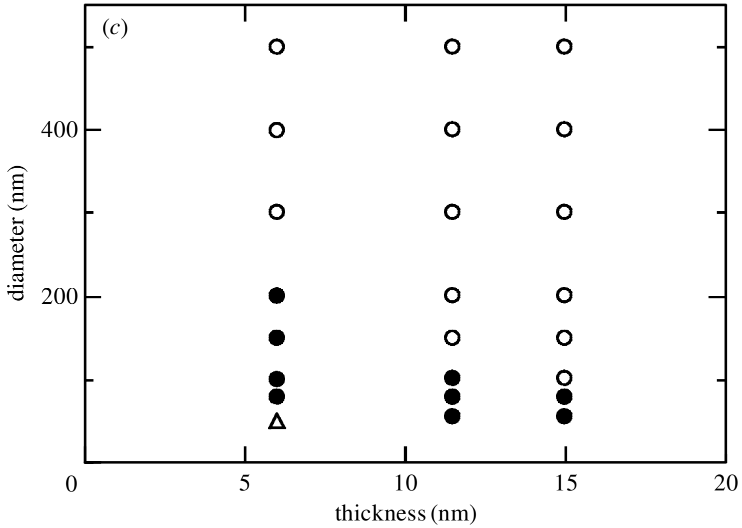

Figure 7. (*Cont.*) (*c*) A phase diagram giving the magnetic state (○, vortex; ●, single domain; △, superparamagnetic) for different diameter and thickness circular nanomagnets.

or exchange stiffness as it is correctly known, opposes any action that prevents the spins being parallel to each other.

Brown understood the perpetual competition that exists between demagnetization and exchange and it led him to what has become known as 'Brown's fundamental theorem'. Brown realized that, in large magnets, demagnetization will prove to be the stronger competitor because of the large surface area of the poles. Conversely, as a magnet is reduced in size, there should come a point at which exchange will gain the upper hand. Very small magnets cannot, therefore, ever be demagnetized, and must adopt the so-called single-domain state. The importance of this point for magnetic data storage cannot be understated. Nanometre-scale magnets (for this is the length-scale on which the transition occurs) are the ideal data-storage device, for they do not lose their memory. It has only been in the last few years that nanotechnology has reached a sufficient state of maturity to allow researchers to begin making and testing these tiny magnets. These are first steps into the field of *quantum engineering*, in which nanotechnology is used to make devices small enough that they access the quantum world (the Pauli exclusion principle and exchange interaction in this case).

Figure 7 shows an experiment that we recently performed (Cowburn *et al.* 2000) demonstrating Brown's fundamental theorem. Using the nanotechnology techniques described above, we made a number of circular nanomagnets. The diameter was in the range 55–500 nm and the thickness was between 6 nm and 15 nm. Figure 7*a*, *b* shows how the hysteresis loops of one of the larger nanomagnets is very different from that of one of the smaller nanomagnets. A 300 nm diameter magnet (figure 7*a*) does not display two different values of magnetization under zero field, and, hence, does not possess a memory function. This is because of demagnetization, which causes a vortex to form at remanence. Conversely, the 100 nm diameter magnet (figure 7*b*) can have either fully positive or fully negative magnetization at zero field, and so is an excellent memory device. Figure 7*c* shows a phase diagram of many different diameters and thicknesses of circular nanomagnet, all obtained directly from experiments, and tells

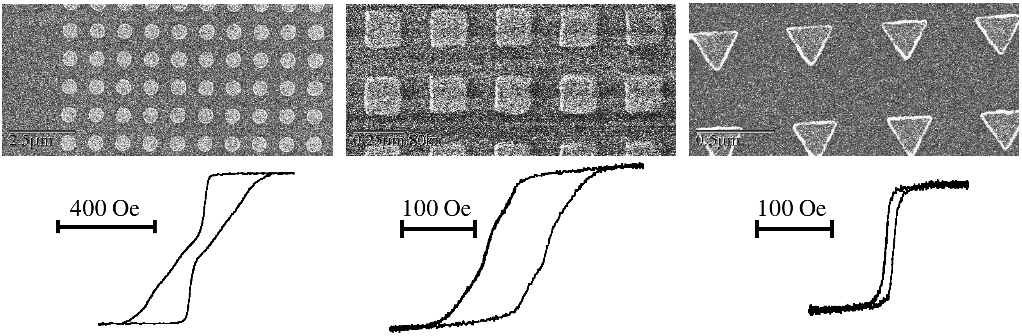

Figure 8. Controlling magnetic properties by nanostructuring: hysteresis loops measured from circular, square and triangular nanomagnets of size 300 nm.

us which sizes and thicknesses are suitable for data storage. As predicted by Brown, only the smaller nanomagnets attain the very useful single-domain state.

The application of nanotechnology to magnetism has so far been presented in the context of one very specific application, namely making new data-storage media. The issue is, however, much wider than this. The latter half of the 20th century has seen the invention of hundreds of different magnetic materials. Given that there are only nine or so commonly used magnetic elements, nearly all of these materials have been made by alloying the nine magnetic elements with each other and with non-magnetic elements. Nanotechnology promises a new generation of artificial magnetic materials. Each nanomagnet is analogous to a giant artificial atom, and one is now free to build new materials, giant atom by giant atom. Figure 8 shows an example of this, where the hysteresis loops of samples of material built from hundreds of circular nanomagnets, square nanomagnets and triangular nanomagnets are compared (Cowburn *et al*. 1998*a*, *b*). Although they are all made from the same material (permalloy) and are of the same thickness (2.5 nm), they all behave very differently and would each be suited to a very different purpose. It is as if we had a wide range of new alloys at our disposal. Magnetic nanotechnology will allow designers to specify the precise magnetic properties they require for the new millennium's magnetic-memory technology and receive a sample of a new material made of artificial giant atoms that possesses just the desired properties.

The quantized magnetic disk (QMD) is one example of a new hard-disk media created by nanotechnology (Chou *et al*. 1994; White *et al*. 1997). Currently only available in research laboratories, the QMD is made of millions of artificially created magnetic pillars. These pillars do not grow by chance, but each is individually placed by the material designer. It is expected that QMD, or patterned media as it is also known, will be used in future generations of hard-disk drives.

Nanotechnology will almost certainly impact the read–write heads in hard-disk drives. A more immediate improvement that can be made is described in § 4, but, ultimately, the single-atom resolution of the scanned probe may be used for reading and writing with ultimate resolution. One possible configuration for this is shown in figure 9, where a pointed nanomagnet is mounted on the end of a scanned probe

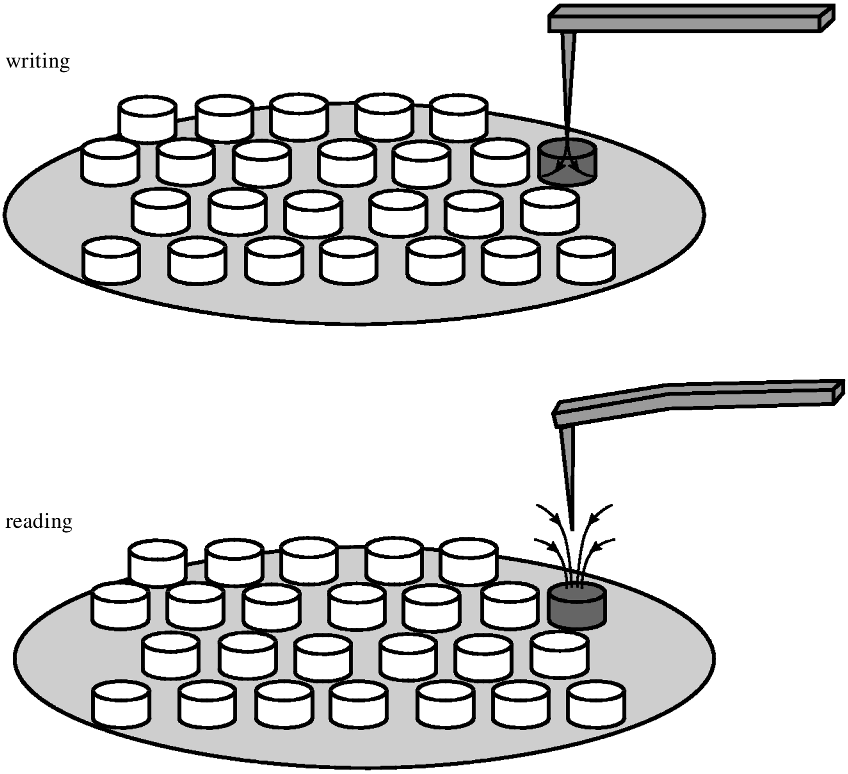

Figure 9. The future of hard-disk data storage. Writing and reading a QMD with a magnetic scanned probe.

cantilever. If the tip is lowered towards the hard-disk surface (in this case assumed to be a QMD), then the magnetic fields coming from the tip will write a data bit onto the disk. If the tip is withdrawn slightly, then the attraction between the tip and the written bit bends the cantilever in a way that can be used to read the bit. The picture of conventionally written hard-disk bits shown in figure 4*b* was obtained in just this way.

4. Spintronics

Spintronics, or magnetoelectronics as it is also known, came about thanks to a very important discovery made in 1988 by the research team of Albert Fert in Paris. Fert asked the simple question: how does the electrical resistance of two individual magnetic films change in a magnetic field when they are placed on top of each other? To their surprise, they found that instead of the tiny variation in resistance that usually arises when a magnetic field is applied to a magnetic material (called anisotropic magnetoresistance (AMR)), they obtained an *enormous* change. The new effect was graphically named giant magnetoresistance, or simply GMR (although an even larger effect has since been discovered in a family of materials called perovskites, which, for lack of superlatives, has had to be called colossal magnetoresistance!). The record for the largest room-temperature resistance change due to GMR is currently held by cobalt–copper multilayers and stands at an enormous 65%. A vitally important

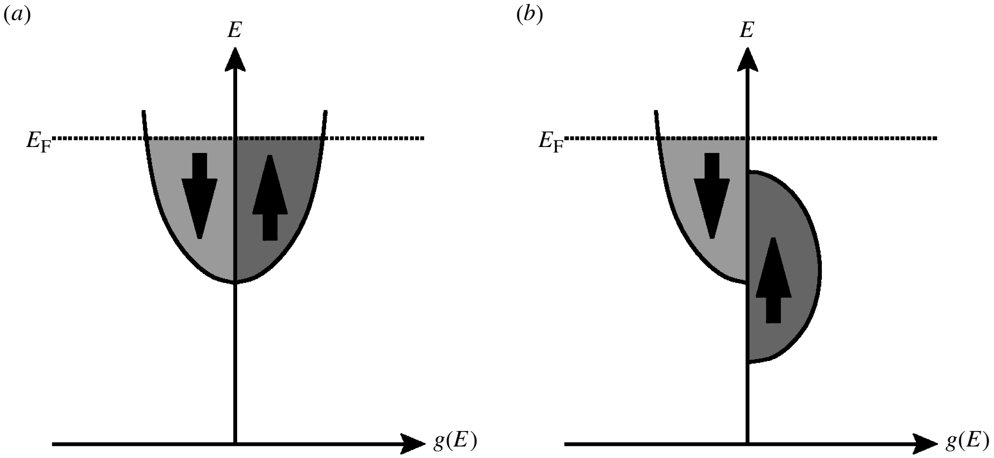

Figure 10. Band diagrams of (*a*) a non-magnetic metal and (*b*) a magnetic metal.

principle had been demonstrated: electronic circuits and magnetic materials are very natural relations. Hence the field of spintronics (magnetic spin plus electronics) was born (Prinz 1999).

Physicists like to describe electrical conduction by a *band diagram* (see figure 10), which is really just a picture of the shape of the bottle that holds a metal's electrons. In a non-magnetic metal, the band containing the spin-up electrons (shorthand for electrons with their spin pointing in an 'upward' direction) is identical to that containing the spin-down electrons. Both types of electron are, therefore, involved in electrical conduction. In a magnetic metal, however, *band splitting* occurs, which means that the band for one of the spin directions is raised in energy a little, and the band for the other spin direction is lowered slightly. Now only electrons at the very top of the band (at the level marked E_F, the Fermi energy) contribute to electrical conduction. Hence, in the case of a magnetic material, all (or most) of the conduction is performed by electrons with their spin in the same direction.

This alone would not cause the enormous effects observed in GMR. To create these, it is necessary to pass current from one piece of magnetic material into another, usually via an intermediary non-magnetic layer (as is shown schematically in figure 11). The electrons leaving the top layer all have their spins pointing in the same direction, say spin-up. GMR arises because, when these electrons try to pass into the second piece of magnetic material, they will only be able to do so if they can find space at the Fermi energy of a band *of the same spin direction*: spin-up electrons can only go into a spin-up bottle. If the two pieces of material are magnetized in the same direction (figure 11*a*), then there is no problem and good conduction occurs. If, however, the pieces of material are magnetized in opposite directions (figure 11*b*), then the first piece can only provide spin-up electrons, whereas the second piece can only accept spin-down electrons because only spin-down states exist at the Fermi level. Hence, there is very little current flow, and a high electrical resistance develops. The device that results is called a *spin valve*, because it is like a water valve. The magnetization directions act as taps for the electric current flow. The spin valve can serve as a very sensitive magnetic-field sensor and also as a memory cell.

The consequences of the GMR effect lie both in the present and the future. The

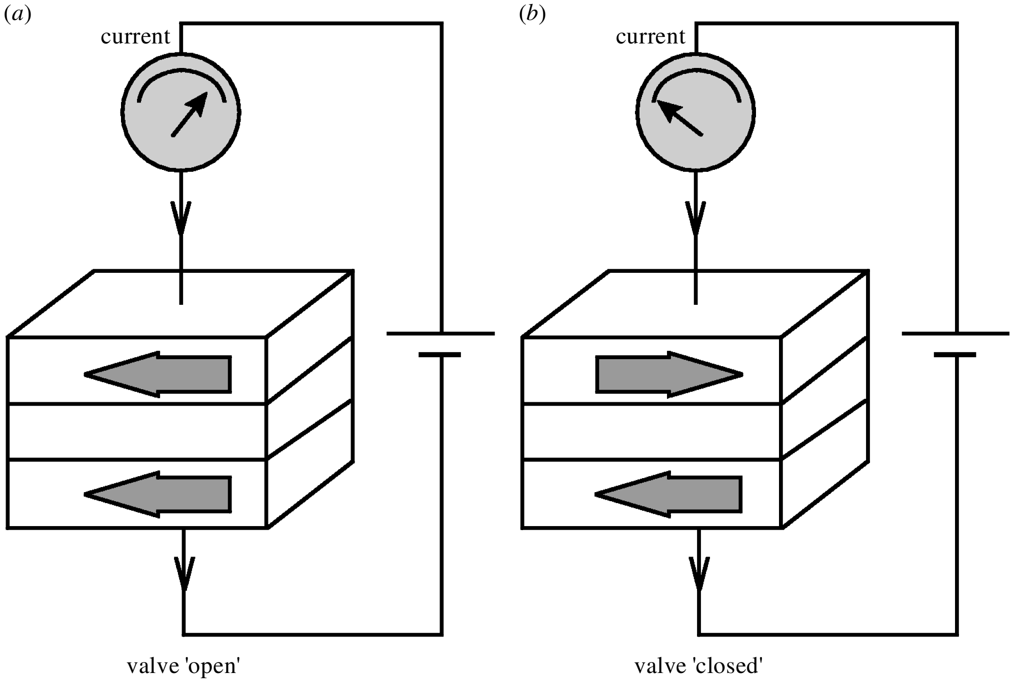

Figure 11. The GMR spin valve. (*a*) Parallel alignment of the two magnetic layers, leading to a low resistance; (*b*) anti-parallel alignment, leading to a high resistance.

currently used AMR heads in hard-disk drives are now being replaced with GMR heads. Using a GMR spin valve as a field sensor allows a sufficiently strong electrical signal to be obtained from correspondingly weaker magnetic fields, and, hence, from much smaller bits.

The future impact of GMR will come from new generations of magnetic memory devices. The most promising of these is the magnetic random access memory chip (MRAM). Currently being developed by companies such as Honeywell and IBM, MRAM seeks to replace electronic memory chips found in computers with magnetic chips. Each bit would be stored in a nanometre-scale piece of magnetic material, and then GMR, or a related effect, would be used to read the data back in electronic form. Figure 12 shows a typical MRAM architecture. A mesh of wires (bit and word lines) forms the addressable area within the chip, with one memory cell capable of storing a 1 or a 0 sitting at each intersection. A current pulse down a single wire creates a magnetic field that alone is insufficient to change the magnetization direction within each storage element. It is only when two fields coincide at one particular bit line–word line vertex that sufficient field is applied to the bit to switch it. In this way, a single bit can be addressed in a large array. Read out is achieved by measuring the resistance of the element. MRAM offers many advantages over conventional memory. Firstly, it could achieve very high storage densities, because only one magnetic element is required per data bit. Semiconductor memory, conversely, requires either more than one transistor per memory cell (SRAM) or a large capacitor per cell (DRAM). Secondly, MRAM is non-volatile (retains its memory when powered down), and so computers would not need lengthy hard-disk reboots every

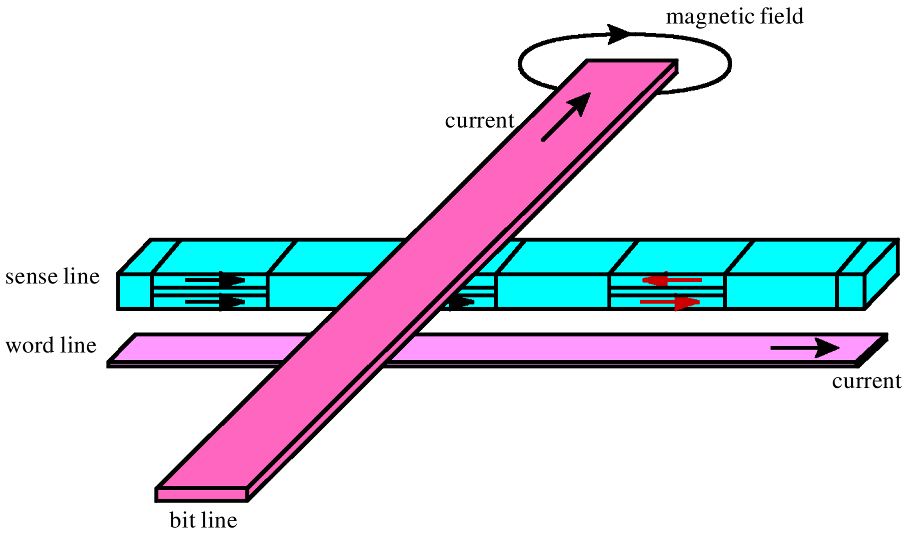

Figure 12. A typical MRAM architecture showing three storage bits.

time they are switched on. The rapidly growing portable computer market desperately needs non-volatile memory that can store information without consuming any power. MRAM is potentially very fast, operating, in principle, on a sub-nanosecond time-scale. A final advantage is that it is not susceptible to radiation damage in the way that semiconductor memory is, and so would be excellent for military and space applications.

MRAM is now established as a future technology and so it is most probable that third-millennium computers will use magnetic memory. Less certain at this time is how much further the conversion of electronics to spintronics will go. Just as hard-disk data storage density has been expanding at an exponential rate, so has microprocessor power and memory-chip capacity. The magnetic transistor has already been demonstrated, and our laboratory, among others, is currently working on magnetic logic gates. A wholesale replacement of electronics by spintronics might be one way in the future to allow continued shrinkage of components.

5. The way ahead

The hard-disk storage industry currently shows no signs of slowing down. Current media will be pushed for a few more years, and so, early in the new millennium, we should be purchasing hard-disk drives with capacities of *ca.* 60 Gb (today's limit is 27 Gb) that work in essentially the same way as today's drives. The AMR heads currently in use will soon be entirely replaced by GMR heads. A further increase in sensitivity of the head might be possible by using CMR materials, but that depends upon development work currently being carried out. After this, today's continuous media will be replaced by patterned media. This might be in the style of the QMD, using a high-throughput production method (contact printing or interferometric optical lithography, both currently under development), or it might involve self-assembly (a chemical process that produces a large number of neatly arranged identical particles). In the first instance, I believe that conventional heads will be used with the patterned media and that just as several grains are currently used to store one bit, so several nanostructures will be covered by a single bit. As head tracking becomes more accurate, this will eventually be reduced to one bit per nanostructure. In this limit, new heads will be needed. These might be scanned-probe heads, or, possibly, spintronic sensors themselves made from state of the art nanotechnology. I believe the ultimate storage density of this technology to be *ca.* 400 Gbits in^{-2} (corresponding to 20 nm size bits), which compares with today's density of 10 Gbits in^{-2}. A typical desktop hard drive would, in this case, be able to store *ca.* 1000 Gb. Two extrapolation lines have been added to figure 2*b* showing when we could expect this limit to be reached. Depending on whether one believes that the current aggressive compound growth rate of 60% per annum (steeper line) could be sustained or that the gentler growth rate that characterized 1960–1990 is more realistic for the future, then these ultimate limit products should be in the shops sometime between 2008 and 2017.

A breakthrough in X-ray optics in the new millennium could change the way the hard drive operates. Magneto-optical (MO) drives currently work very successfully, using a pulse of laser light to read and write the magnetic disk. Current laser sources do not allow much scope for future development, however, because the minimum bit size is determined by the wavelength of the laser light (*ca.* 500 nm). If a portable and cheap X-ray laser is discovered in the new millennium, then MO may yet re-enter the race.

High-performance refrigeration may become an accepted part of data storage, as the superparamagnetic problem can be eliminated by working at low temperatures. Other developments in computing—such as single-electron transistors, molecular electronics and superconducting switches—would also benefit from refrigeration. The issues preventing high-performance desk-top computers from being sold with a liquid helium refrigeration unit are currently mainly institutional (the industry is not familiar or comfortable with the idea) and not technical.

I predict that the currently exponential growth of hard-disk storage density will eventually flatten off. When this happens (around 2014 according to my predictions), the industry will undergo a phase change. Progress will no longer be measured by storage density, but by diversity. The range of hard-drive products currently available is actually very small given that the industry has annual revenues in excess of $30 billion. Once it is no longer economically viable to push the data density any further, attention will switch to making a broader range of products, each adapted to a specific situation. Some drives will be very small, others very fast and others very robust.

After the industrial phase change, the hard-drive industry will begin to decline. I say this because I assume that a new data storage paradigm will be found in the new millennium. Magnetic storage has been, and still will be, phenomenally successful, but all things have their day. The real problem with current magnetic storage is that it is intrinsically two dimensional. Data are only stored on the surface of a disk, which is very wasteful. Three-dimensional data storage is, in my opinion, the future for data storage in the third millennium. There is some chance that it will still use a magnetic principle, but that is not yet certain. The advantages of going to three dimensions can be demonstrated by a simple calculation. My prediction of 400 Gbits in^{-2} as the ultimate maximum was made assuming one data bit every 40 nm on the surface of a disk. If, now, the same size bit is used, but this time inside a three-dimensional solid, then the equivalent storage density is 260 000 000 Gbits in^{-3}! In the ultimate limit of one atom representing one data bit, a piece of material the size of a sugar lump could store 10^{25} bits, or 10 000 000 000 000 000 Gbits. The numbers are unimaginable! The big problem facing this idea is one of addressing: how exactly does one read and write to the atom in the middle of the block? Although a mighty technological challenge, I believe that a solution could be found within the first century of the third millennium. The answer may well involve using electromagnetic waves, especially if a cheap X-ray laser is available.

I firmly believe that magnetic RAM chips will become a reality very soon in the new millennium. The technical difficulties involved are, in my opinion, surmountable. The biggest difficulties are institutional. A semiconductor fabrication plant is an unbelievably expensive purchase. Today's price tag is around $2 billion and rising. The inertia arising from this scale of investment creates a very understandable conservatism within the industry, making paradigm shifts difficult to make, because any progress has to be made by evolution and not revolution. Perhaps this will be less of a problem in the new millennium. Should the semiconductor industry itself undergo a major change in structure (perhaps moving to many small manufacturing sites due to advances in fabrication technology and a market requirement to diversify product range), then perhaps new ideas such as MRAM will have more latitude to develop.

One of the interesting aspects of the development of the hard-disk industry on the one hand and of MRAM on the other is that the two technologies are on converging paths. Conventionally, hard-disk storage is cheaper (per bit) than RAM because the data storage sites do not need to be predefined. If the new millennium sees a move towards patterned media, and semiconductor RAM chips become magnetic RAM chips, then the only difference between the two will be the addressing method: MRAM will use a mesh of wires, whereas the QMD will use a flying head. The problems facing the flying head will become increasingly acute, however, as the bits become smaller. A point may soon come when it is simpler to overlay the QMD with a mesh of wires and address the individual bits that way. At that point, hard-disk and RAM technologies will have merged.

The third millennium may see new applications of magnetism in addition to data storage. Many science undergraduate coffee rooms buzz with talk of a quantum computer. Currently only a theoretical construct, the quantum computer uses the uncertainty of the quantum world to perform many calculations simultaneously. Unfortunately, the theory is still far ahead of the practice. Because the theory is so general, it is currently not even known which branch of science to use to implement the idea. Magnetism may yet prove to be the most favourable. Very small nanomagnets, such as those used in the experiment of figure 7 and smaller, could perhaps be used, although probably only at very low temperature. The major difficulty facing the quantum computer is how to control errors. The problem becomes less acute the smaller one goes in size, and so nanotechnology will certainly be needed. One idea currently being considered is to use the very small magnetic moment contained in the *nucleus* of the atom. This is not so far fetched as it might seem, for already hospital MRI scanners do precisely this. A related but separate development that is currently in its infancy uses a scanned probe cantilever, such as the one shown in figure 5*b*, to do MRI scans (or nuclear magnetic resonance (NMR) as it is more correctly termed) across very small objects. The goal of this work is to visualize the individual atoms in a molecule such as a protein using magnetic imaging. This would be the ultimate biochemical analysis tool.

There is one issue in data-storage technology in the third millennium that remains to be addressed, and, in many ways, it is the most important. Why do we need so much storage capacity? A modest hard drive can today store all of the *Encyclopaedia Brittanica* with plenty of space left over, and how many PC owners believe they will read even all of that in a lifetime? In my opinion, we need more storage capacity *not* to do more of the same, but to do *different* things. I tentatively suggest that if storage capacities grow by many orders of magnitude (as envisaged by three-dimensional storage), then two major changes will come about in society as a result.

The first is that it will change the nature of information. Traditionally, we think of information as being words. Until recently, computer data storage could only really handle text. Graphics could be stored, but they filled the disk space very rapidly. If the available space was much much greater, then all kinds of different forms could be used for communicating: pictures, sounds, animation, etc. Just as poetry is better for expressing love than it is for teaching somebody how to use a video recorder, we will be free to choose the information form most suited to the content. If this is combined with a widespread use of e-books (a form of portable computer that might replace books in the next millennium), then the nature of information really will change. The currently separate roles of authors, artists, actors and musicians would all merge in the creation of a new pan-form art. This can only happen if we can store it all.

The second change in society that I see coming about as a result of an unlimited ability to store information is, perhaps, the most important, and is a change in the way we understand history. The second millennium witnessed an enormous breaking down of *spatial* barriers. For most of the inhabitants of the 11th century, the neighbouring village was a far away place, seldom visited. Most inhabitants of the early 21st century will have visited most of the major cities in their country, many of the countries on their continent and several countries in the world. In contrast, the *temporal* barriers that divide one generation from another remain, and, if anything, are more pronounced than they were a millennium ago. Mass data storage is not a time machine. It is, however, a temporal telescope, allowing us to see clearly into the past. Suppose that every detail of life in the 21st century were able to be stored. Would not the inhabitants of the 22nd century understand their past, and, hence, to some extent, themselves, more clearly? Just as science was the most successful academic discipline of the latter half of the second millennium, giving us great insights into the natural world and an ability to control it for our benefit, perhaps history will become the dominant academic discipline of the third millennium. Mass information storage will provide an abundance of primary historical sources, allowing the historian to delve deeply into the workings of the human world, just as the development of the experimental scientific method provided the unbounded number of primary scientific sources that allows science to work so well. The development of the *form* of information will be essential in this. Anyone who has wasted an afternoon drowning in Internet detritus looking for a single reference will know that more stored data is not necessarily more accessible information. I would hope that progress in information technology during the third millennium, building on the second millennium's first attempt at virtual reality, would allow the historian to really spend a day in the life of a person of (future) antiquity. Advanced bio-interfaces would allow that person's thoughts to be rethought, their feelings refelt and their full world view re-experienced. To understand history correctly is to understand the influences that have made us the people we are today, and, therefore, to be able to weigh our opinions more freely and reasonably. If future progress in magnetic storage technology can allow mankind to do that, then magnetism does indeed have an attractive future.

This work would not have been possible without the assistance of my colleagues Professor Mark Welland, Dr Kunle Adeyeye and Denis Koltsov and the support of St John's College, Cambridge, and The Royal Society.

References

Chou, S. Y., Wei, M. S., Krauss, P. R. & Fischer, P. B. 1994 Single domain magnetic pillar array of 35 nm diameter and 65 Gbit/in^2 density for ultrahigh density quantum magnetic storage. *J. Appl. Phys.* **76**, 6673.

Cowburn, R. P., Koltsov, D. K., Adeyeye, A. O. & Welland, M. E. 1998*a* Probing sub-micron nanomagnets by magneto-optics. *Appl. Phys. Lett.* **73**, 3947.

Cowburn, R. P., Adeyeye, A. O. & Welland, M. E. 1998*b* Configurational anisotropy in nanomagnets. *Phys. Rev. Lett.* **81**, 5414.

Cowburn, R. P., Adeyeye, A. O., Welland, M. E. & Tricker, D. M. 1999 Single domain circular nanomagnets. *Phys. Rev. Lett.* **83**, 1042.

Grochowski, E. & Thompson, D. A. 1994 Outlook for maintaining areal density growth in magnetic recording. *IEEE Trans. Mag.* **30**, 3797.

Malhotra, S. S., Lal, B. B., Alex, M. & Russak, M. A. 1997 Effect of track edge erasure and on-track percolation on media noise at high recording density in longitudinal thin film media. *IEEE Trans. Mag.* **33**.

Prinz, G. A. 1999 Magnetoelectronics. *Science* **283**, 330.

Verschuur, G. L. 1993 *Hidden attraction: the history and mystery of magnetism.* Oxford University Press.

White, R. L., New, R. M. H. & Pease, R. F. W. 1997 Patterned media: a viable route to 50 Gbit/in^2 and up for magnetic recording? *IEEE Trans. Mag.* **33**, 990.

AUTHOR PROFILE

R. P. Cowburn

Russell Cowburn was born in Newcastle upon Tyne. After a year working in the defence electronics industry, he arrived in Cambridge in 1990 to study Natural Sciences. He graduated in 1993 with first class honours and immediately began a PhD at the Cavendish Physics Laboratory, Cambridge. His research into the magnetism of films of atomic thickness led him to spend a year working at the CNRS, Paris. In 1997 he was elected to a research fellowship at St John's College, Cambridge, to carry out research (part funded by a Royal Society research grant) into magnetism and nanotechnology. Aged 28, he has published over 25 papers and has been an invited speaker at five international conferences. Married and a committed Christian, his interests include hill walking, Brahms and P. G. Wodehouse.