

Serial: Computer Memory

In the previous installment, we explained why electronics must be continuously improved and where its current limits lie. We also explained the concept of *spin* and how its transport gives rise to spin current. In this installment, we will present one real-world spintronic[1] component that we encounter every day and that has substantially changed data storage capabilities — the *hard disk* — and another component that combines the properties of a high-capacity memory with the speed of RAM — *magnetoresistive RAM (MRAM)*.

Before introducing these two technologies, let us consider what we actually expect from such a memory. We have three requirements:

1. **Small binary structures.** We need a technology that uses physical properties with two very precisely defined states, representing our ones and zeros. Furthermore, we want to be able to create well-defined spatial structures that are very small. This allows us to concentrate large amounts of information in a small area.

2. **Long-term stability.** For a memory to be truly useful, it must be able to retain information for a long time. For example, standard hard disk memories retain information for several decades.

3. **Non-invasive reading.** The stored information must be readable without being affected by the reading process.

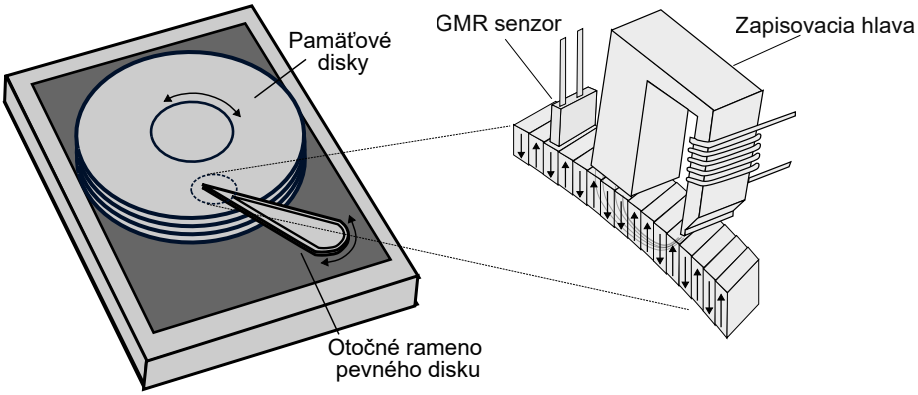


Figure 1: Illustration of a hard disk. On the left is a drawing of the entire hard disk; on the right is a detailed cross-section of a hard disk platter and read head.

Let us look at a typical hard disk, shown in Figure 1. As its name suggests, the core of the technology is a disk or a set of disks covered with a special magnetic layer. A read/write head

[1] The hard disk itself is based on magnetism, but it contains a spintronic component: a GMR sensor.

is mounted on a movable arm next to the disk. By mechanically moving the arm and rotating the disk, it is possible to access any location on the disk with high precision. When multiple disks are stacked on top of one another, the arms are also stacked, and the control unit simply determines which head is currently reading or writing. The read head consists of a GMR sensor, which we will discuss later. Writing is performed using a special coil that focuses the magnetic field and thereby switches the magnetic moments in the magnetic layer.

Bits, Bytes, and Structures

In conventional HDDs, data is written to thin ferromagnetic layers. The reason is that, in ferromagnets, the magnetic moments of individual atoms are aligned in the same direction. As a result, they produce a strong magnetic field even without an external field, which can be easily detected by the read head. Moreover, this configuration of moments in ferromagnets is entirely spontaneous (it is maintained without the influence of an external magnetic field) and is caused by the so-called exchange interaction, which can effectively be understood as the action of a very strong magnetic field on the order of thousands of tesla. Microscopically, the exchange interaction has a quantum-mechanical origin and is beyond the scope of this series. What matters is that it allows a given state to be retained for a long time without additional energy input, since this state is natural for ferromagnets.

Real materials, however, also have certain preferred directions in which their magnetization tends to settle. These are described by the so-called magnetic anisotropy. From a microscopic perspective, there are several sources of anisotropy. For example, shape anisotropy favors certain directions based on the shape of the object, while magnetocrystalline anisotropy favors a particular equilibrium orientation due to the symmetry of the crystal. The preferred magnetization directions must reflect the symmetry of the crystal structure of real materials, and this is precisely what allows us to represent "ones and zeros" in a ferromagnetic layer.

To work with these contributions, we need to establish the appropriate mathematical framework. In this series, we will describe anisotropy using the energy density ε. Mathematically, anisotropy can be thought of as an additional effective magnetic field that modifies the effects of the exchange interaction. We begin with the simplest possible model of magnetic anisotropy, namely *uniaxial magnetic anisotropy*. It creates two preferred magnetization directions along the corresponding axis. Together, these give us the free-energy functional $\varepsilon(\theta)$ in the form

$$\varepsilon(\theta) = K_\mathrm{u} \sin^2\theta - \mu_0 M_\mathrm{s} H \cos(\theta - \theta_H)\,, \tag{1}$$

where θ is the angle determining the direction of magnetization, K_u is a constant describing the magnetic anisotropy, H is the magnetic field strength, directed at an angle θ_H, μ_0 is the vacuum permeability, and M_s is the saturation magnetization. An example of the energy functional is shown on the left in Figure 2.

A simple example of calculating a hysteresis loop can be demonstrated for the case $\theta_H = = 0°$. The model used can already predict a realistic representation of information using the two states of a single bit, corresponding to two magnetization orientations. Ferromagnets are characterized by so-called magnetization hysteresis, which means that the magnetic state depends not only on the current conditions but also on the previous states of the system. For example, consider a ferromagnet magnetized so that all its moments point upward, and suppose we begin applying a magnetic field directed downward. The ferromagnet retains its state until the external field exerts a force sufficient to flip the moments. In our energy functional $\varepsilon(\theta)$, this

corresponds to a state in which the first derivative of the system's energy is zero (the system is in an equilibrium state, which is not necessarily stable) and its second derivative is negative (the system is unstable, and even small perturbations cause it to switch to the nearest energy minimum). Although the hysteresis loop can be calculated analytically in simple cases, we will use the provided script `hysteresis.m` [2], which numerically checks these conditions, finds the minima, and finally plots the hysteresis loop. To run it successfully, we need to fill in the energy functional (lines 26–27) and its first and second derivatives (lines 30–31 and 34–35).

A possible result of such a simulation for an illustrative energy functional is shown on the right in Figure 2. With zero applied magnetic field, the magnetization is in its original position A ($\theta = 0\,\mathrm{deg} = 0°$). Another minimum is at point D, where $\theta = 180\,\mathrm{deg} = 180°$, corresponding to the second stable position. When a magnetic field of positive polarity is applied, the energy functional at point A (or point $\theta = 0\,\mathrm{deg}$) begins to increase. The system is in a non-optimal state, but not enough for it to jump to the second, global minimum. The limit is point B, where the minimum essentially disappears, and even small perturbations switch the system to the second minimum. As H increases further, the system jumps to point C ($\theta = = 180\,\mathrm{deg}$), where the magnetization flips. If we now switch off the magnetic field, the system returns to point D, since this is still the local energy minimum closest to $\theta = 180\,\mathrm{deg}$. Points A, B, C, and D are also marked in the figure on the right, where we can see the formation of hysteresis. Applying the same procedure for the opposite polarity of H gives us the complete loop.

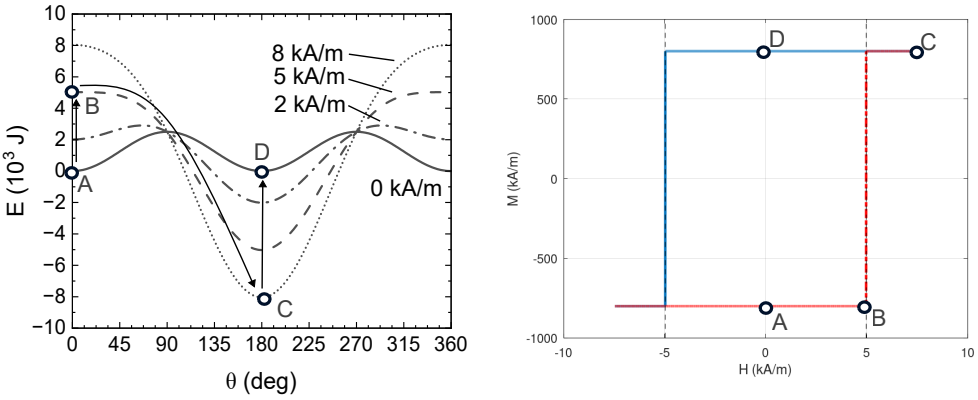


Figure 2: Left: energy functional[3] of a ferromagnetic layer for $M_\mathrm{s} = 800\,\mathrm{kA{\cdot}m^{-1}}$, $K_\mathrm{u} = 2\,500\,\mathrm{J}$, $\theta_H = 0°$, and various applied magnetic fields H ranging from $0\,\mathrm{kA{\cdot}m^{-1}}$ to $8\,\mathrm{kA{\cdot}m^{-1}}$. Right: magnetization M as a function of the applied magnetic field H.

It is important to emphasize that this is the simplest possible model. To describe the real behavior of more complex materials, we need to account for several additional effects (strain, crystal structure effects, etc.), which can simply be added to Equation (1). In the

[2] `https://drive.google.com/file/d/14lFLXKfB1QJweUY-MvRvVVEELoShf7aU/view?usp=sharing`

[3] The graph uses the total energy rather than the energy density, but the distinction is irrelevant for our purposes. The constant K_u is likewise used in the context of energy.

next installment of the series, you will be able to simulate hysteresis loops yourselves for the diluted magnetic semiconductor GaMnAs, which was once considered a candidate material for spintronic devices.

From a practical perspective, the ideal hysteresis loop would be as narrow as possible, so that writing requires as little energy as possible. However, the situation is not that simple. So far, we have completely neglected fluctuations of the moments, which arise due to the nonzero ambient temperature. The magnetic anisotropy determines how long the information is retained (that is, how long it takes for the magnetic moments to spontaneously rotate due to thermal fluctuations). In practice, it is therefore necessary to choose the right material, one that retains information over a long period while also requiring little energy to switch the magnetic moments.

But how can we create small regions for storing individual bits? Under certain conditions, a ferromagnet naturally divides into magnetic domains — regions in which the magnetization points approximately in the same direction. Dividing into domains reduces the magnetic field energy around the object. However, creating domain walls requires energy, so the resulting domain structure depends on the size, shape, and material of the ferromagnet. Sufficiently small particles can remain single-domain.

Natural domains, however, do not by themselves provide a regular and stable arrangement of memory cells. Hard disks therefore use granular magnetic layers consisting of tiny ferromagnetic grains separated by nonmagnetic boundaries. These suppress the exchange coupling between neighboring grains. A single recorded bit generally comprises several grains whose magnetizations are aligned in the same direction during writing. Their anisotropy then helps preserve this state even after the writing field is removed.

Sensors

We have found a way to write and store bits. But how do we detect them? We can measure a magnetic field using, for example, a coil, with the motion of the rotating disk platter producing small changes in the magnetic flux. Such a sensor, however, is not sufficiently sensitive. Instead, methods based on changes in electrical resistance caused by a magnetic field are used. This phenomenon is called magnetoresistance, specifically giant magnetoresistance (GMR).

Such a sensor can be complicated, but the basic idea can be understood using a model with two ferromagnetic layers F separated by a nonmagnetic layer N. The purpose of this arrangement is to produce different scattering rates for electrons at the interface between the ferromagnet and the nonmagnetic spacer. If an electron's spin is parallel to the magnetization, it is scattered less than an electron with the opposite spin. An interesting situation arises, however, when we construct the F|N|F "sandwich" shown in Figure 3. We fix the magnetization of one layer and leave the other free to rotate. A magnetic field from the "bits" (or magnetic domains) can then reverse the magnetization of the sensor's free magnetic layer. In the parallel configuration (the magnetizations of both layers point in the same direction), electrons with only one spin orientation (opposite to the magnetization) are strongly scattered, resulting in low resistance. Conversely, in the antiparallel configuration (the magnetizations of the layers point in opposite directions), both spin-up and spin-down electrons are strongly scattered (each type in a different layer), resulting in a sharp increase in resistance. Quantitatively, these changes can reach hundreds of percent, making it possible to detect even small magnetic domains and thus increase hard disk capacity.

To conclude this subsection, we should add that even more sensitive sensors currently exist, for example, those based on tunneling magnetoresistance (TMR), which operate on a similar but physically more complex principle. Curious readers can find a detailed discussion of this topic in the specialist literature[4].

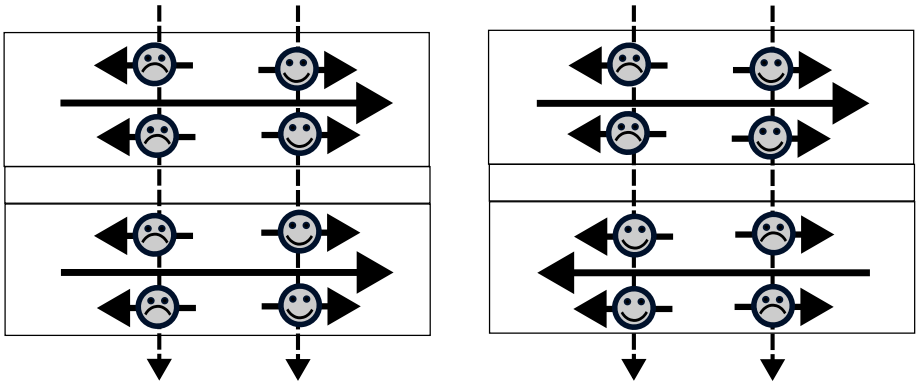


Figure 3: Schematic of a GMR sensor. On the left is the parallel (low-resistance) state; on the right is the antiparallel (high-resistance) state. The sad smiley face indicates strongly scattered electrons.

The Future of Hard Disks

Although the operating principle of hard disks presented here may seem simple, it is the result of decades of development that have made it possible to achieve the capacities available today. And their development is far from over. It is interesting to look at current trends and challenges in this field. These can be found in so-called roadmaps (development plans). For hard disks, these represent the industry's collective outlook and address how much data we want to be able to store on a disk in the future and what technologies we need to develop to achieve this. Roadmaps are published by disk manufacturers, their suppliers, and researchers, either jointly through professional associations or by individual companies in the form of their own plans. The goal of roadmaps is to help coordinate research, manufacturing, and investment so that individual components are ready at approximately the same time. For example, while the areal recording density was estimated at 1 Tb per square inch in 2018, it is expected to exceed 10 Tb per square inch by 2035, and companies are adjusting their activities accordingly.

So what are the current trends? We have essentially already mentioned them above, but the goals are even more ambitious — we want to "bring the bits" even closer together and reduce the energy required to write them. Alloys of iron and platinum in the form of grains are currently being developed for the manufacture of hard disk platters. This ensures that individual grains are separated by nonmagnetic gaps, which reduce the magnetic interaction between them. This makes it possible to write a bit into a well-defined number of grains and reduces the area occupied by each bit. If we look at the 2022 roadmap, we find that the artificial

[4]COEY, J. M. D. Magnetic tunnel junctions. In: *Magnetism and Magnetic Materials*. Cambridge: Cambridge University Press, 2009. ISBN 978-0-521-81614-4.

creation of so-called magnetic islands is being considered. In such modified islands, a bit can be written into a single, well-defined ferromagnetic domain.

Can we also reduce the energy required to rewrite a bit? One possibility is to add another mechanism to the writing process. For example, as we will show in subsequent installments, a high-frequency electromagnetic wave can set the magnetization into motion; once the magnetization is rotating, it is easier to switch. This phenomenon is used by the so-called MAMR (Microwave-Assisted Magnetic Recording) technology. Even more interesting from an application perspective is the so-called HAMR (Heat-Assisted Magnetic Recording) technology, in which a thin laser beam briefly increases the temperature of the ferromagnet, thereby reducing its magnetic anisotropy and making writing easier. Once implemented, these two technologies could increase the areal recording density to approximately 10 Tb per square inch in the near future.

MRAM

Another technology that we will briefly introduce is magnetoresistive random-access memory (MRAM). As mentioned last time, the advantage of hard disks is their ability to reliably retain information for a long time. However, they are slow. RAM, on the other hand, is fast, but compared with a hard disk, it stores only a fraction of the amount of information, and only while continuously powered. So why not combine the properties of magnetic materials, in which the exchange interaction holds spins together, with fast access to information? This is precisely the question behind MRAM technology.

MRAM uses the orientation of magnetization to store information, which it reads electrically through magnetoresistance. The main motivation behind its design is to combine fast reading and writing with high endurance against repeated rewriting and nonvolatility (the retention of data even after the power supply is disconnected). Unlike DRAM (dynamic RAM), it does not require regular data refreshes and allows inactive memory blocks to be powered down without losing their contents. MRAM thus offers a path toward more energy-efficient embedded memories and, potentially, certain levels of processor cache.

In this installment, we will introduce only the zeroth generation of MRAM, since the technology that improved it will be the subject of the next installment of the series. In the zeroth generation, a grid of nanometer-scale wires was created, with a TMR sensor at each intersection (as shown in Figure 4). Information was written using the magnetic field generated by currents in the addressing wires. Writing occurred only at the intersection of the wires, because the magnetic fields from both wires added together, producing a sufficiently strong resultant field to write the information. These states were read as low or high resistance through tunneling magnetoresistance. The disadvantages were high writing currents and the risk of inadvertently switching neighboring cells exposed to part of the writing field. These problems were solved by spin-orbit torque (SOT) and spin-transfer torque (STT) technologies.

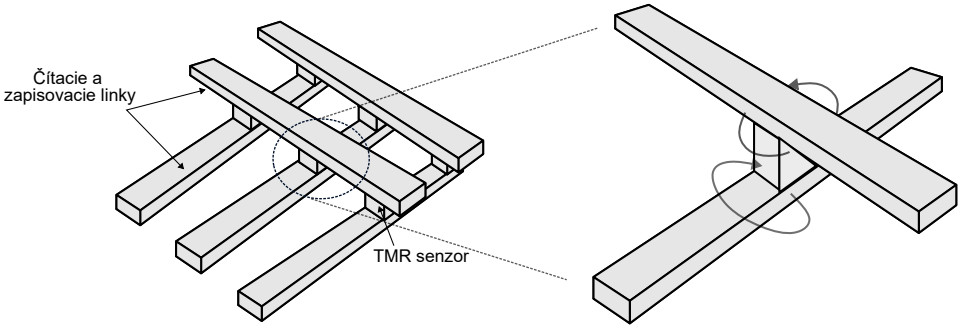


Figure 4: Illustration of field MRAM (an example of zeroth-generation MRAM). TMR sensors are located at the intersections and record the information. Writing is performed by simultaneously applying fields from the read and write lines, as illustrated in the enlarged view on the right.

Peter Kubaščík
peter.kubascik@fykos.org

FYKOS is organized by students of Faculty of Mathematics and Physics of Charles University. It's part of Media Communications and PR Office and is supported by Institute of Theoretical Physics of CUNI MFF, his employees and The Union of Czech Mathematicians and Physicists. The realization of this project was supported by Ministry of Education, Youth and Sports of the Czech Republic.

This work is licensed under Creative Commons Attribution-Share Alike 3.0 Unported.
To view a copy of the license, visit https://creativecommons.org/licenses/by-sa/3.0/.