Astronomy

Instability could explain the Sun's curious cycle

Ellen Zweibel

A phenomenon that affects the magnetic fields of rotating bodies could be involved in recurring changes in the Sun's behaviour, which are related to a periodic flipping of its field. The proposal is a fresh take on this strange effect. **See p.769**

High-energy radiation from the Sun follows an 11-year cycle that has profound consequences for life on Earth. For example, it changes the temperature of Earth's upper atmosphere and affects the behaviour of space-borne instruments, such as communications satellites. This cycle is related to the Sun's magnetic field, but despite decades of observational and theoretical progress, a consistent explanation for many aspects of solar magnetism's most basic features remains elusive. On page 769, Vasil *et al.*[1] add a provocative ingredient to the theoretical mix that could prove key to unravelling this astrophysical enigma.

A brief description of the solar cycle is in order[2]. The Sun's magnetic field is similar in some ways to Earth's magnetic field, which is a giant dipole that is oriented roughly along its axis of rotation. But the Sun's field is skewed, so that the lines have a 'toroidal' component, which runs parallel to the Sun's equator (Fig. 1). Dark spots are visible where the toroidal lines emerge from the surface and the concentration of these sunspots changes gradually over a period of 11 years, as the toroidal component migrates from mid to equatorial latitudes. This cycle coincides with a reversal of the polarity of the Sun's magnetic field.

The field's 'poloidal' component runs parallel to lines of constant longitude, and its reversal lags the changes in the toroidal component by a quarter of a cycle. The magnetic helicity, which measures the magnitude and direction of field-line twist, is positive in the Sun's southern hemisphere and negative in its northern hemisphere.

It is widely accepted that this cycle of magnetic activity is sustained by flows in the solar interior and at the surface, which form a magnetohydrodynamic dynamo (a kind of generator that uses energy from the movement of an electrically conducting fluid to form and maintain a magnetic field). Because the electrical conductivity is very high in the Sun's interior, the gas dictates the shape of the field almost everywhere, and the fluid flow amplifies the field by stretching it.

One example of this flow-induced field amplification is a mechanism called the Zeldovich dynamo[3]. Imagine that the magnetic field takes the shape of an elastic band. The surrounding fluid stretches and twists the loop until it resembles a figure of eight, which is then folded to form two loops. When these loops merge (again, through flow-induced effects), the resulting field has the same extent as the original field, but twice its strength. The field, in turn, influences the fluid flow: curved field lines exert a tension force, which induces a flow that makes the field lines shorten or straighten. This back reaction of the field on the flow results in dynamos being inherently constrained in terms of the strength of the field.

Three types of flow occur inside the Sun, all of which can be probed with the tools of helioseismology. The first is thermal convection, which transports heat to the surface from a layer in the interior known as the thermal convection zone, much as a liquid heated from below is driven to boil. This convection occurs in discrete 'cells', and the Sun's rotation imparts a twist to these cells as they rise and fall, similarly to how cyclones form on Earth where there are strong currents of rising air. The second occurs because the rotation rate varies with depth and latitude — this type of flow is known as differential rotation, and it is universally seen in liquids and gases in rotating bodies. The third type is poloidal, or meridional, circulation[4], in which fluid moves at the surface from the equator towards the poles.

Early models[5] of the solar cycle were based on a mechanism known as a dynamo loop, in which the toroidal field was generated from the poloidal field through stretching caused by differential rotation, and the poloidal field was produced from the toroidal field through effects caused by turbulent convection. Subsequent models incorporated three more ingredients: the tachocline, a layer of gas that is subject to strong shear forces at the base of the thermal convection zone; the transport of magnetic flux by meridional circulation;

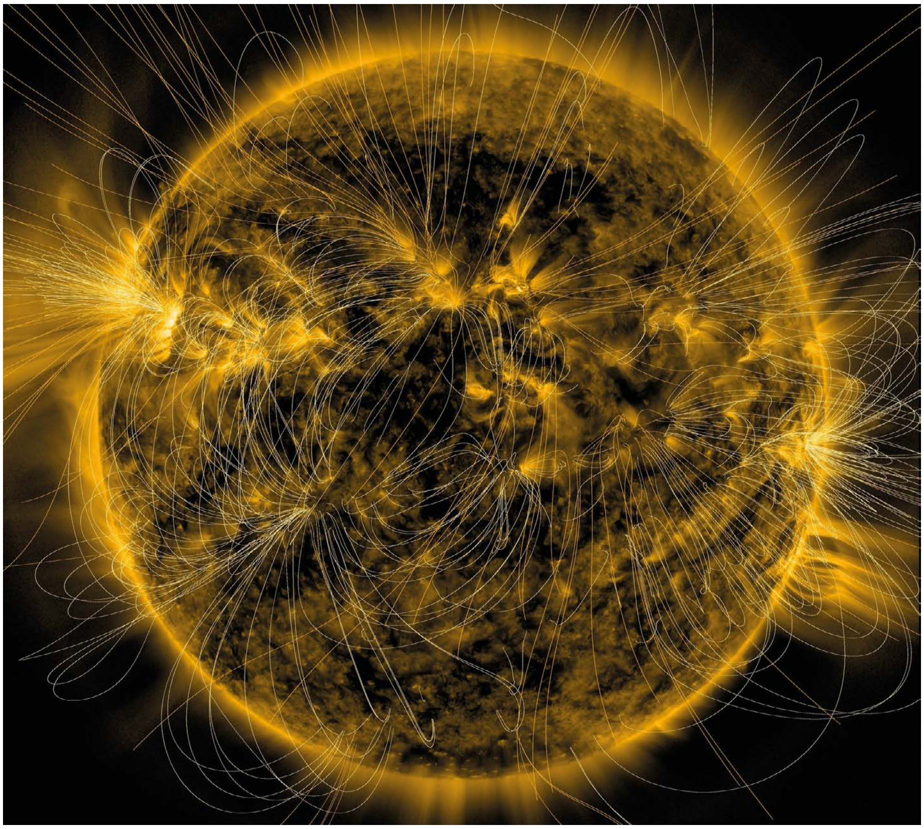
NASA/SDO/AIA/LMSAL

Figure 1 | Ultraviolet image of the Sun. Brightness indicates a strong surface magnetic field. The overlaid lines represent an extrapolation of the magnetic field above the surface.

and another layer known as the near-surface shear layer – although its importance has been largely discounted[6].

At the same time, advances in computing made it possible to perform simulations of convective, differentially rotating, magnetized shells in the hope of replicating the solar cycle[7]. However, no model – be it numerical or a stripped-down theoretical one – has yet explained the length or amplitude of the solar cycle without a fine-tuning of parameters. And if future simulations succeed in capturing cycles with the observed features, it would still be necessary to understand the mechanisms at play to apply the model to other stars, and to understand the Sun's behaviour over long timescales.

There is one kinematic feature that has not yet been mentioned: the solar cycle is linked to fluctuations in the rate at which the surface rotates, and these are known as torsional oscillations. Although a relationship between the torsional oscillations and the solar magnetic field has been discussed before[8], no model has yet made the connection that Vasil *et al.* now propose – namely, that both are manifestations of the same underlying phenomenon, known as the magnetorotational instability (MRI), which arises when an electrically conducting fluid in a magnetic field spins faster near its centre of rotation than it does farther away.

The MRI is generally considered to be the main mechanism through which compact objects, such as black holes, accrete matter from a rapidly rotating disc: the instability transports angular momentum outwards[9], allowing matter to move inwards. This occurs because the magnetic field associated with the object generates tension that makes fluid displaced inwards continue to fall towards the object. Turbulence arising from the MRI has also been shown to amplify magnetic fields, thereby making the MRI a dynamo. This provides a natural way of limiting the amplitude of the instability: if the field is sufficiently strong, magnetic tension becomes the dominant force and it reins in amplification of the field. This exemplifies the interplay between fluid flow and curved field lines.

Vasil *et al.* show, through simplified analysis and numerical simulations, that an initially poloidal magnetic field in the near-surface shear layer can be amplified by the MRI. The radial shear forces that are produced by the MRI generate a toroidal field, accompanied by rotational velocity fluctuations, which the authors identify as the torsional oscillations. Importantly, their picture is consistent with the magnetic helicity being positive in the Sun's southern hemisphere and negative in its northern hemisphere. They also find that the rate at which the instability grows is comparable to the period of the solar cycle, offering a tantalizing clue to the otherwise obscure timing of the cycle. And the instability has a natural way of saturating – by enhancing the magnetic field and by modifying the structure of the near-surface shear layer so as to stabilize it.

The authors' model is highly simplified, especially in its treatment of thermal convection, and the existence of the near-surface shear layer is not explained. However, Vasil and colleagues' initial results are intriguing. They could well furnish an interpretative framework for more-elaborate models, and they are sure to inspire future studies.

Ellen Zweibel is in the Departments of Astronomy and Physics, University of Wisconsin, Madison, Wisconsin 53706, USA.
e-mail: zweibel@astro.wisc.edu

1. Vasil, G. M. *et al. Nature* **629**, 769–772 (2024).
2. Hathaway, D. H. *Living Rev. Sol. Phys.* **12**, 4 (2015).
3. Zeldovich, I. B., Ruzmaikin, A. A. & Sokolov, D. D. *Magnetic Fields in Astrophysics* (Gordon & Breach, 1990).
4. Hanasoge, S. M. *Living Rev. Sol. Phys.* **19**, 3 (2022).
5. Charbonneau, P. *Living Rev. Sol. Phys.* **17**, 4 (2020).
6. Dikpati, M., Corbard, T., Thompson, M. J. & Gilman, P. A. *Astrophys. J.* **575**, L41–L45 (2002).
7. Charbonneau, P. & Sokoloff, D. *Space Sci. Rev.* **219**, 35 (2023).
8. Howard, R. & Labonte, B. J. *Astrophys. J.* **239**, L33–L36 (1980).
9. Balbus, S. A. & Hawley, J. F. *Astrophys. J.* **376**, 214–222 (1991).

The author declares no competing interests.

In retrospect

40 years since genomic imprinting was discovered

Anne C. Ferguson-Smith & Marisa S. Bartolomei

Some genes carry an 'imprint' on either the maternal or the paternal copy, which determines whether or not that copy is expressed. This 1984 discovery changed how scientists think about gene regulation and inheritance.

In mammals, the expression of certain genes depends on which parent they were inherited from. For most genes in a cell, both copies are turned either on or off. But for a small subset of genes in mammals, one copy is on and the other off. For some of these genes, it is the maternal copy that is on; for others, it is the paternally inherited copy. This remarkable phenomenon, known as genomic imprinting, was discovered 40 years ago in landmark embryo-manipulation experiments[1–3] reported by Surani, Barton and Norris in *Nature*[1,3] and by McGrath and Solter in *Cell*[2].

The experiments found that mouse embryos with two sets of chromosomes (diploid) failed to complete development if both sets of chromosomes were derived either from the female parent (bimaternal) or from the male parent (bipaternal). These papers proved that both parental genomes are essential for normal mammalian development. Importantly, they showed that maternally and paternally inherited chromosomes are not functionally equivalent, and that each copy of the genome carries distinguishing 'imprints' that are set during the formation of the parents' eggs or sperm (gametes). These imprints are now known to be biochemical changes to DNA, known as epigenetic modifications, that mark the genes as being either on or off after fertilization.

Before the 1984 experiments, scientists had shown that diploid embryos that were created by manipulating eggs so that the embryos lacked paternally derived chromosomes (that is, they were parthenogenetic) failed to develop to term[4]. However, these studies could not rule out other factors as contributing to the non-viability, including deficiencies in the egg's cytoplasm. This, combined with irreproducible findings that diploid embryos lacking paternal chromosomes could yield viable and fertile adult female mice[5], meant that the 1984 embryological investigations were conducted at a time of heightened technical concerns, and of outcomes that remained open to interpretation.

Newly fertilized mammalian eggs contain two pronuclei: one with a set of chromosomes from the egg; the other with a set from the sperm. The discovery of genomic imprinting depended on the success of sophisticated 'reconstitution' experiments in which the maternal or paternal pronucleus (or both) was isolated from a donor embryo and fused to a recipient one-cell embryo with one or both of its pronuclei physically removed. This proved to be the most effective way to generate diploid embryos in which it was possible to control the origin of the two parental genomes, the